

```
1 NUMBER OF SEQ ID NOS: 24
2 SOFTWARE: Patentin version 3.0
3 SEQ ID NO 2
4 LENGTH: 640
5 TYPE: PRT
6 ORGANISM: Bacillus sp.
7 FEATURE:
8 NAME/KEY: misc_feature
9 LOCATION: (3)..(3)
10 OTHER INFORMATION: Xaa is any amino acid
11 NAME/KEY: misc_feature
12 LOCATION: (24)..(24)
13 OTHER INFORMATION: Xaa is any amino acid
14 NAME/KEY: misc_feature
15 LOCATION: (30)..(30)
16 OTHER INFORMATION: Xaa is any amino acid
17 NAME/KEY: misc_feature
18 LOCATION: (33)..(33)
19 OTHER INFORMATION: Xaa is any amino acid
20 NAME/KEY: misc_feature
21 LOCATION: (47)..(47)
22 OTHER INFORMATION: Xaa is any amino acid
23 NAME/KEY: misc_feature
24 LOCATION: (48)..(48)
25 OTHER INFORMATION: Xaa is any amino acid
26 NAME/KEY: misc_feature
27 LOCATION: (54)..(54)
28 OTHER INFORMATION: Xaa is any amino acid
29 NAME/KEY: misc_feature
30 LOCATION: (71)..(71)
31 OTHER INFORMATION: Xaa is any amino acid
32 NAME/KEY: misc_feature
33 LOCATION: (75)..(75)
34 OTHER INFORMATION: Xaa is any amino acid
35 NAME/KEY: misc_feature
36 LOCATION: (90)..(90)
37 OTHER INFORMATION: Xaa is any amino acid
38 NAME/KEY: misc_feature
39 LOCATION: (103)..(103)
40 OTHER INFORMATION: Xaa is any amino acid
41 NAME/KEY: misc_feature
42 LOCATION: (106)..(106)
43 OTHER INFORMATION: Xaa is any amino acid
44 NAME/KEY: misc_feature
45 LOCATION: (129)..(129)
46 OTHER INFORMATION: Xaa is any amino acid
47 NAME/KEY: misc_feature
48 LOCATION: (131)..(131)
49 OTHER INFORMATION: Xaa is any amino acid
50 NAME/KEY: misc_feature
51 LOCATION: (132)..(132)
52 OTHER INFORMATION: Xaa is any amino acid
53 NAME/KEY: misc_feature
54 LOCATION: (133)..(133)
55 OTHER INFORMATION: Xaa is any amino acid
56 NAME/KEY: misc_feature
57 LOCATION: (134)..(134)
58 OTHER INFORMATION: Xaa is any amino acid
59 NAME/KEY: misc_feature
60 LOCATION: (147)..(147)
61 OTHER INFORMATION: Xaa is any amino acid
62 NAME/KEY: misc_feature
63 LOCATION: (149)..(149)
64 OTHER INFORMATION: Xaa is any amino acid
65 NAME/KEY: misc_feature
66 LOCATION: (161)..(161)
67 OTHER INFORMATION: Xaa is any amino acid
68 NAME/KEY: misc_feature
69 LOCATION: (166)..(166)
70 OTHER INFORMATION: Xaa is any amino acid
71 NAME/KEY: misc_feature
72 LOCATION: (173)..(173)
73 OTHER INFORMATION: Xaa is any amino acid
```

```
1 NAME/KEY: misc_feature
2 LOCATION: (184)..(184)
3 OTHER INFORMATION: Xaa is any amino acid
4 NAME/KEY: misc_feature
5 LOCATION: (188)..(188)
6 OTHER INFORMATION: Xaa is any amino acid
7 NAME/KEY: misc_feature
8 LOCATION: (189)..(189)
9 OTHER INFORMATION: Xaa is any amino acid
10 NAME/KEY: misc_feature
11 LOCATION: (190)..(190)
12 OTHER INFORMATION: Xaa is any amino acid
13 NAME/KEY: misc_feature
14 LOCATION: (195)..(195)
15 OTHER INFORMATION: Xaa is any amino acid
16 NAME/KEY: misc_feature
17 LOCATION: (287)..(287)
18 OTHER INFORMATION: Xaa is any amino acid
19 NAME/KEY: misc_feature
20 LOCATION: (307)..(307)
21 OTHER INFORMATION: Xaa is any amino acid
22 NAME/KEY: misc_feature
23 LOCATION: (325)..(325)
24 OTHER INFORMATION: Xaa is any amino acid
25 NAME/KEY: misc_feature
26 LOCATION: (370)..(370)
27 OTHER INFORMATION: Xaa is any amino acid
28 NAME/KEY: misc_feature
29 LOCATION: (432)..(432)
30 OTHER INFORMATION: Xaa is any amino acid
31 NAME/KEY: misc_feature
32 LOCATION: (502)..(502)
33 OTHER INFORMATION: Xaa is any amino acid
34 NAME/KEY: misc_feature
35 LOCATION: (532)..(532)
36 OTHER INFORMATION: Xaa is any amino acid
37 NAME/KEY: misc_feature
38 LOCATION: (542)..(542)
39 OTHER INFORMATION: Xaa is any amino acid
40 NAME/KEY: misc_feature
41 LOCATION: (585)..(585)
42 OTHER INFORMATION: Xaa is any amino acid
43 NAME/KEY: misc_feature
44 LOCATION: (592)..(592)
45 OTHER INFORMATION: Xaa is any amino acid
46 NAME/KEY: misc_feature
47 LOCATION: (593)..(593)
48 OTHER INFORMATION: Xaa is any amino acid
49 NAME/KEY: misc_feature
50 LOCATION: (595)..(595)
51 OTHER INFORMATION: Xaa is any amino acid
52 NAME/KEY: misc_feature
53 LOCATION: (596)..(596)
54 OTHER INFORMATION: Xaa is any amino acid
55 NAME/KEY: misc_feature
56 LOCATION: (597)..(597)
57 OTHER INFORMATION: Xaa is any amino acid
58 NAME/KEY: misc_feature
59 LOCATION: (612)..(612)
60 OTHER INFORMATION: Xaa is any amino acid
61 NAME/KEY: misc_feature
62 LOCATION: (633)..(633)
63 OTHER INFORMATION: Xaa is any amino acid
64 US-09-509-814A-2
```

Query Match

Best Local Similarity 85.9%; Score 1941.5; DB 4; Length 640;
Matches 374; Conservative 21; Mismatches 38; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAONNYGLVGGQGVAVADTGLDTRGNDSSMEHAFRGKITALVALGRIN 60
DB 207 NDVARGIVKADVAONNYGLVGGQGVAVADTGLDTRGNDSSMEHAFRGKITALVALGRIN 266


```

Db      1 NDVARGIVKADVAQSSVGLYGQGIYAVADTGLDGTGRNDSMHEARFGKITLALGRN 60
Qy      61 NANDTNGHGHVAGSVLGNATNKMAPQANLVFOSIMDSGGGLGSLPSNLQTLFSQAFS 120
Db      61 NANDTNGHGHVAGSVLGNATNKMAPQANLVFOSIMDSGGGLGSLPSNLQTLFSQAFS 120
Qy      121 AGARIHNSWGAAVNGAYTTDSRVDDYVRKNDMTILFAAGNEPNGGTTISAPGTAKNAI 180
Db      121 AGARIHNSWGAAVNGAYTTDSRVDDYVRKNDMTILFAAGNEPNGGTTISAPGTAKNAI 180
Qy      181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLPDSSF 240
Db      181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLPDSSF 240
Qy      241 WANHDSKYAWGTSMAPIVAGNVAQLREHFYKNGITPKPSLLKALILAGADVGLGY 300
Db      241 WANHDSKYAWGTSMAPIVAGNVAQLREHFYKNGITPKPSLLKALILAGADVGLGY 300
Qy      301 PNGNGGGRVTLDKSLNVAAYNNESSALSTSOKATYTFETATAGKPKLSLWSDAPASTTA 360
Db      301 PNGNGGGRVTLDKSLNVAAYNNESSALSTSOKATYTFETATAGKPKLSLWSDAPASTTA 360
Qy      361 SVTLVNDLDELITAPNGTRVYGNDFSAFPDNNWDGRNNEVNFINSPOSGTYTIEVOAYN 420
Db      361 SVTLVNDLDELITAPNGTRVYGNDFSAFPDNNWDGRNNEVNFINSPOSGTYTIEVOAYN 420
Qy      421 VPVGPOKPSLAIYN 434
Db      421 VPVGPOKPSLAIYN 434

```

RESULT 2

AA17089 standard; protein; 639 AA.

```

XX      AA17089;
AC      20-MAR-2003 (revised)
DT      21-JUL-1999 (first entry)
XX      DE Bacillus alkaline protease.
XX      KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX      washing composition; oxidizing agent.
OS      Bacillus sp.
XX      PN W09918218-A1.
XX      PD 15-APR-1999.
XX      PF 07-OCT-1998; 98WC-JP004528.
XX      PR 07-OCT-1997; 97JP-00274570.
XX      (KAOS ) KAO CORP.
XX      PI Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
XX      PI Shikata S, Nomura M,
XX      DR WPI; 1999-287736/27.
XX      DR N-PSDB; AAK37277.
XX      PT Alkali protease from Bacillus used in washing powders.
XX      PS Disclosure; Page 53-58; 71pp; Japanese.
CC      The invention relates to alkaline proteases produced by strains of
CC      Bacillus; the proteases ability to digest casein is not inhibited by
CC      oleic acid and they have a high stability to oxidizing agents. The
CC      alkaline protease of the invention has the following properties: (a) it
CC      is active over the pH range 4-13 and has at least 80% of its optimum

```

activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d) its ability to digest casein is not inhibited by oleic acid; (e) it has molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be used as enzymes in washing compositions for use in automatic dishwashers and for washing clothes. The stability to oxidizing agents allows the enzyme to be an effective component of washing compositions including bleach. The present sequence represents an alkaline protease. (updated on 20-MAR-2003 to correct DR field.)

Sequence 639 AA:

Query Match 99.6%; Score 2237; DB 2; Length 639;
 Best Local Similarity 99.8%; Pred. No. 6.6e-155;
 Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy      1 NDVARGIVKADVAQSSVGLYGQGIYAVADTGLDGTGRNDSMHEARFGKITLALGRN 60
Db      206 NDVARGIVKADVAQSSVGLYGQGIYAVADTGLDGTGRNDSMHEARFGKITLALGRN 265
Qy      61 NANDTNGHGHVAGSVLGNATNKMAPQANLVFOSIMDSGGGLGSLPSNLQTLFSQAFS 120
Db      266 NANDTNGHGHVAGSVLGNATNKMAPQANLVFOSIMDSGGGLGSLPSNLQTLFSQAFS 325
Qy      121 AGARIHNSWGAAVNGAYTTDSRVDDYVRKNDMTILFAAGNEPNGGTTISAPGTAKNAI 180
Db      326 AGARIHNSWGAAVNGAYTTDSRVDDYVRKNDMTILFAAGNERPNGGTTISAPGTAKNAI 385
Qy      181 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLPDSSF 240
Db      386 TVGATENLRPSFGSYADNINHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLPDSSF 445
Qy      241 WANHDSKYAWGTSMAPIVAGNVAQLREHFYKNGITPKPSLLKALILAGADVGLGY 300
Db      446 WANHDSKYAWGTSMAPIVAGNVAQLREHFYKNGITPKPSLLKALILAGADVGLGY 505
Qy      301 PNGNGGGRVTLDKSLNVAAYNNESSALSTSOKATYTFETATAGKPKLSLWSDAPASTTA 360
Db      506 PNGNGGGRVTLDKSLNVAAYNNESSALSTSOKATYTFETATAGKPKLSLWSDAPASTTA 565
Qy      361 SVTLVNDLDELITAPNGTRVYGNDFSAFPDNNWDGRNNEVNFINSPOSGTYTIEVOAYN 420
Db      566 SVTLVNDLDELITAPNGTRVYGNDFSAFPDNNWDGRNNEVNFINSPOSGTYTIEVOAYN 625
Qy      421 VPVGPOKPSLAIYN 434
Db      626 VPVGPOKPSLAIYN 639

```

RESULT 3

AA17091 standard; protein; 640 AA.

```

XX      AA17091;
AC      20-MAR-2003 (revised)
DT      21-JUL-1999 (first entry)
XX      DE Bacillus alkaline protease.
XX      KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX      washing composition; oxidizing agent.
OS      Bacillus sp.
XX      PN W09918218-A1.
XX      PD 15-APR-1999.
XX      PF 07-OCT-1998; 98WC-JP004528.
XX      PR 07-OCT-1997; 97JP-00274570.
XX

```

PA (KAOS) KAO CORP.
 XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
 PI Shikata S, Nomura M;
 XX MPI: 1999-287726/27.
 DR N-PSDB; AAK37279.
 XX
 XX Alkali protease from *Bacillus* used in washing powders.
 PS Disclosure; Page 63-68; 71pp; Japanese.
 XX
 CC The invention relates to alkaline proteases produced by strains of
 CC *Bacillus*. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidizing agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the range pH 6-11; (b) after 30 minutes at 40 deg C it is
 CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidizing agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease. (Updated
 CC on 20-MAR-2003 to correct DR field.)
 CC
 SQ Sequence 640 AA:
 Query Match 97.9%; Score 2197; DB 2; Length 640;
 Best Local Similarity 96.8%; Pred. No. 5,6e-152;
 Matches 440; Conservative 13; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVADVAGSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITLVALGRTN 60
 DB 207 NDVARGIVADVAGSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITLVALGRTN 266
 QY 61 NANDTNGHGHVAGSVLGNATKKGMAPOANLVFQSIMDSGGGLPSNLQTLFSGQAFS 120
 DB 267 NANDTNGHGHVAGSVLGNATKKGMAPOANLVFQSIMDSGGGLPSNLQTLFSGQAFS 326
 QY 121 AGARHTNSGGAAVNGAYTTDSRNVDYRKNDMTLLFAAGNEGPGGTISAPGTAKNAI 180
 DB 327 AGARHTNSGGAAVNGAYTTDSRNVDYRKNDMTLLFAAGNEGPGGTISAPGTAKNAI 386
 QY 181 TVGATEENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTYIISARSSLAPDSSF 240
 DB 387 TVGATEENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTYIISARSSLAPDSSF 446
 QY 241 WANHDSKTYAVNGGTSNATPIVAGNVAQLREHFVKRNGITPKPSLLKALINAGADVGLGY 300
 DB 447 WANHDSKTYAVNGGTSNATPIVAGNVAQLREHFVKRNGITPKPSLLKALINAGADVGLGY 506
 QY 301 PNGNGMGWVTLDKSLNVAAYVNESSALSTSQKATYTFATAGKPLKISLVMSDAPASTTA 360
 DB 507 PNGNGMGWVTLDKSLNVAAYVNESSALSTSQKATYTFATAGKPLKISLVMSDAPASTTA 566
 QY 361 SVTLVNDLDELVTAPNGTRYVGNDSAPFDNNMDGNNVENVFINSPOSGTYTIEVQAYN 420
 DB 567 SVTLVNDLDELVTAPNGTRYVGNDSAPFDNNMDGNNVENVFINSPOSGTYTIEVQAYN 626
 QY 421 VVVGPNQFSLATVN 434
 DB 627 VVVGPNQFSLATVN 640

RESULT 4
 AAM50080
 ID AAM50080 standard; protein; 434 AA.
 XX AAM50080;
 AC
 XX
 DT 12-AUG-2002 (first entry)

XX
 DE *Bacillus* sp KSM-KP43 alkaline protease protein fragment.
 XX
 KM Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX
 OS *Bacillus* sp.
 XX
 PN EPI209233-A2.
 XX
 PD 29-MAY-2002.
 XX
 XX 22-NOV-2001; 2001EP-00127851.
 PF
 XX 22-NOV-2000; 2000JP-00355166.
 PR
 XX 12-APR-2001; 2001JP-00114048.
 XX
 PA (KAOS) KAO CORP.
 PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N,
 PI Okuda M, Saeki K;
 XX MPI: 2002-437518/47.
 DR
 PT New modified alkaline proteases useful in detergent compositions.
 XX
 PS Claim 1; Page 10-11; 25pp; English.
 XX
 CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency %
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 28%). This
 CC sequence represents a fragment of the alkaline protease KP43 from
 CC *Bacillus* sp strain KSM-KP43 which is used to create the modified protease
 CC represented in AAM50090
 CC
 SQ Sequence 434 AA:
 Query Match 97.6%; Score 2191; DB 5; Length 434;
 Best Local Similarity 96.5%; Pred. No. 9.1e-152;
 Matches 419; Conservative 13; Mismatches 2; Indels 0; Gaps 0;
 QY 1 NDVARGIVADVAGSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITLVALGRTN 60
 DB 1 NDVARGIVADVAGSSYGLYGQGIYAVADTGLDTRGNDSSMHEAFRGKITLVALGRTN 60
 QY 61 NANDTNGHGHVAGSVLGNATKKGMAPOANLVFQSIMDSGGGLPSNLQTLFSGQAFS 120
 DB 61 NANDTNGHGHVAGSVLGNATKKGMAPOANLVFQSIMDSGGGLPSNLQTLFSGQAFS 120
 QY 121 AGARHTNSGGAAVNGAYTTDSRNVDYRKNDMTLLFAAGNEGPGGTISAPGTAKNAI 180
 DB 121 AGARHTNSGGAAVNGAYTTDSRNVDYRKNDMTLLFAAGNEGPGGTISAPGTAKNAI 180
 QY 181 TVGATEENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTYIISARSSLAPDSSF 240
 DB 181 TVGATEENLRPSFGSYADNINHVAQFSSRGPTKDGRIKPDVMAFGTYIISARSSLAPDSSF 240
 QY 241 WANHDSKTYAVNGGTSNATPIVAGNVAQLREHFVKRNGITPKPSLLKALINAGADVGLGY 300
 DB 241 WANHDSKTYAVNGGTSNATPIVAGNVAQLREHFVKRNGITPKPSLLKALINAGADVGLGY 300
 QY 301 PNGNGMGWVTLDKSLNVAAYVNESSALSTSQKATYTFATAGKPLKISLVMSDAPASTTA 360
 DB 301 PNGNGMGWVTLDKSLNVAAYVNESSALSTSQKATYTFATAGKPLKISLVMSDAPASTTA 360
 QY 361 SVTLVNDLDELVTAPNGTRYVGNDSAPFDNNMDGNNVENVFINSPOSGTYTIEVQAYN 420
 DB 361 SVTLVNDLDELVTAPNGTRYVGNDSAPFDNNMDGNNVENVFINSPOSGTYTIEVQAYN 420
 QY 421 VVVGPNQFSLATVN 434
 DB 421 VVVGPNQFSLATVN 434

XX Sequence 639 AA; Score 2155; DB 2; Length 639;
Query Match 96.0%; Best Local Similarity 96.3%; Pred. No. 6.5e-149; Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
1 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMHEAFRGKITLALYALGRTN 60
206 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMHEAFRGKITLALYALGRTN 265
QY 61 NNDNTNGHGTTHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGSLPNTLQTLFSGQAFS 120
Db 266 NNDNTNGHGTTHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGSLPNTLQTLFSGQAFS 325
QY 121 AGARHTNSWGAANVNGAYTTDSRNVDYRKNDMTLLFAAGNEGPNCGTISAPGTAKNAI 180
Db 326 AGARHTNSWGAANVNGAYTTDSRNVDYRKNDMTLLFAAGNEGPNCGTISAPGTAKNAI 385
QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTYLLSARSSILAPDSSF 240
Db 386 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTYLLSARSSILAPDSSF 445
QY 241 WANHSKRYAVMGSTSMATPIVAGNVAQLREHFVKRGTIPKPSLLKALILAGAADVGLGY 300
Db 446 WANHSKRYAVMGSTSMATPIVAGNVAQLREHFVKRGTIPKPSLLKALILAGAADVGLGY 505
QY 301 PNGNQGWGRVTLDKSLNVAAYVNESSALSTSQKATYFTATAGKPLKISLVMSDAPASTTA 360
Db 506 PNGNQGWGRVTLDKSLNVAAYVNESSALSTSQKATYFTATAGKPLKISLVMSDAPASTTA 565
QY 361 SVTLVNDLDLVITAPNGTRYVGNDFSAFPDMMDGRRNVNVTINSQSTYITTEVOAYN 420
Db 566 SVTLVNDLDLVITAPNGTRYVGNDFSAFPDMMDGRRNVNVTINSQSTYITTEVOAYN 625
QY 421 VPVGPOXFSLAIVN 434
Db 626 VPVGPOXFSLAIVN 639
RESULT 7
AA17088 standard; protein; 640 AA.
XX
AC AA17088;
XX
DT 20-MAR-2003 (revised)
DT 21-JUL-1999 (first entry)
XX
DE An alkaline protease sequence from Bacillus species.
XX
KM Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
XX washing composition; oxidizing agent.
XX
OS Bacillus sp.
XX
XX Key Location/Qualifiers
FH Misc-difference 1..640
FT /note= "all residues indicated as Xaa are arbitrary amino
FT acids"
XX
XX W09918218-A1.
XX
XX 15-APR-1999.
XX
XX 07-OCT-1998; 98WO-JP004528.
XX
XX 07-OCT-1997; 97JP-00274570.
XX
XX (KAO) KAO CORP.
XX
XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y,
PI Shikata S, Nomura M,
PI

XX WPI: 1999-287736/27.
DR N-PSDB; AAX37278.
XX
FT Alkali protease from Bacillus used in washing powders.
XX
PS Claim 3; Page 50-53; 71pp; Japanese.
XX
CC The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidizing agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidizing agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease of the
CC invention. (Updated on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 640 AA;
Query Match 96.0%; Score 2155; DB 2; Length 640;
Best Local Similarity 96.3%; Pred. No. 6.5e-149; Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
1 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMHEAFRGKITLALYALGRTN 60
207 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMHEAFRGKITLALYALGRTN 266
QY 61 NNDNTNGHGTTHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGSLPNTLQTLFSGQAFS 120
Db 267 NNDNTNGHGTTHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGSLPNTLQTLFSGQAFS 326
QY 121 AGARHTNSWGAANVNGAYTTDSRNVDYRKNDMTLLFAAGNEGPNCGTISAPGTAKNAI 180
Db 327 AGARHTNSWGAANVNGAYTTDSRNVDYRKNDMTLLFAAGNEGPNCGTISAPGTAKNAI 386
QY 181 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTYLLSARSSILAPDSSF 240
Db 387 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTYLLSARSSILAPDSSF 446
QY 241 WANHSKRYAVMGSTSMATPIVAGNVAQLREHFVKRGTIPKPSLLKALILAGAADVGLGY 300
Db 447 WANHSKRYAVMGSTSMATPIVAGNVAQLREHFVKRGTIPKPSLLKALILAGAADVGLGY 506
QY 301 PNGNQGWGRVTLDKSLNVAAYVNESSALSTSQKATYFTATAGKPLKISLVMSDAPASTTA 360
Db 507 PNGNQGWGRVTLDKSLNVAAYVNESSALSTSQKATYFTATAGKPLKISLVMSDAPASTTA 566
QY 361 SVTLVNDLDLVITAPNGTRYVGNDFSAFPDMMDGRRNVNVTINSQSTYITTEVOAYN 420
Db 567 SVTLVNDLDLVITAPNGTRYVGNDFSAFPDMMDGRRNVNVTINSQSTYITTEVOAYN 626
QY 421 VPVGPOXFSLAIVN 434
Db 627 VPVGPOXFSLAIVN 640
RESULT 8
AAMS0085 standard; protein; 434 AA.
XX
AC AAMS0085;
XX
DT 12-AUG-2002 (first entry)
DT
XX
XX Bacillus sp alkaline protease protein A-1 fragment.
DE
XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
KW

```

XX OS Bacillus sp.
XX PN EPI209233-A2.
XX PD 29-MAY-2002.
XX PF 22-NOV-2001; 2001EP-00127851.
XX PR 22-NOV-2000; 2000JP-00355166.
XX PR 12-APR-2001; 2001JP-00114048.
XX PA (KAO) KAO CORP.
XX PI Hatada Y, Ogawa A, Kagayama Y, Sato T, Araki H, Sumitomo N;
XX PI Okuda M, Saeki K;
XX DR WPI; 2002-437518/47.
XX PT New modified alkaline proteases useful in detergent compositions.
XX PS Claim 5; Page 18-19; 25pp; English.
XX CC This invention describes novel Bacillus sp. alkaline proteases useful in
XX CC detergent compositions, especially in laundry, bleaching or automatic
XX CC dishwasher detergents. The novel proteases have an increased detergency &
XX CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX CC sequence represents a fragment of the alkaline protease A-1 from Bacillus
XX CC sp NCIB12289 described in the method of the invention
XX SQ Sequence 434 AA;

Query Match 94.9%; Score 2130; DB 5; Length 434;
Best Local Similarity 92.4%; Pred. No. 2.6e-147;
Matches 401; Conservative 24; Mismatches 9; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMEAFRGKITALTALYALGRTN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMEAFRGKITALTALYALGRTN 60
QY 61 NADNTNGHGHVAGSVLGNATNKMAPQANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 120
DB 61 NADPNHGHTHVAGSVLGNATNKMAPQANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 120
QY 121 AGARIHTNSWGAANVAGYTTDSRVVDYVRKNDMTILFAAGNEGPNQGTISAPGTAKNAI 180
DB 121 AGARIHTNSWGAAPVNGAYTTDSRVVDYVRKNDMTILFAAGNEGPNQGTISAPGTAKNAI 180
QY 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTKGRIPDVMAQPTIYLSARSSLAPDSGF 240
DB 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTKGRIPDVMAQPTIYLSARSSLAPDSGF 240
QY 241 WANDSKKAYMGGSMTPIVAGVNAQLREHFVNRGITPKPSILRAALTAGADVGLGY 300
DB 241 WANDSKKAYMGGSMTPIVAGVNAQLREHFVNRGITPKPSILRAALTAGADVGLGY 300
QY 301 PNGNQGWGRVTLDSKLVNAVYNESSALSTSOKATYFTATAGKPLKISLWSDAPASTTA 360
DB 301 PNGNQGWGRVTLDSKLVNAVYNESSALSTSOKATYFTATAGKPLKISLWSDAPASTTA 360
QY 361 SVTLVNDLDTVITAPNGTRVYGNDFSAFPDNNMDGRNNVENVFINSQSGTYTVEQAYN 420
DB 361 SVTLVNDLDTVITAPNGTRVYGNDFSAFPDNNMDGRNNVENVFINSQSGTYTVEQAYN 420
QY 421 VPVGPNFSLAIYN 434
DB 421 VPVGPNFSLAIYN 434

```

```

AC AAM50086;
XX 12-AUG-2002 (first entry)
DT Bacillus sp alkaline protease protein A-2 fragment.
DE Bacillus sp alkaline protease protein A-2 fragment.
XX DE Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX KM Bacillus sp.
XX OS EPI209233-A2.
XX PN 29-MAY-2002.
XX PD 22-NOV-2001; 2001EP-00127851.
XX PF 22-NOV-2000; 2000JP-00355166.
XX PR 12-APR-2001; 2001JP-00114048.
XX PA (KAO) KAO CORP.
XX PI Hatada Y, Ogawa A, Kagayama Y, Sato T, Araki H, Sumitomo N;
XX PI Okuda M, Saeki K;
XX DR WPI; 2002-437518/47.
XX PT New modified alkaline proteases useful in detergent compositions.
XX PS Claim 5; Page 20-21; 25pp; English.
XX CC This invention describes novel Bacillus sp. alkaline proteases useful in
XX CC detergent compositions, especially in laundry, bleaching or automatic
XX CC dishwasher detergents. The novel proteases have an increased detergency &
XX CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX CC sequence represents a fragment of the alkaline protease A-2 from Bacillus
XX CC sp NCIB12513 described in the method of the invention
XX SQ Sequence 433 AA;

Query Match 94.4%; Score 2118.5; DB 5; Length 433;
Best Local Similarity 93.3%; Pred. No. 1.8e-146;
Matches 405; Conservative 18; Mismatches 10; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMEAFRGKITALTALYALGRTN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGIYVADTGLDTGRNDSMEAFRGKITALTALYALGRTN 60
QY 61 NADNTNGHGHVAGSVLGNATNKMAPQANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 120
DB 61 NADPNHGHTHVAGSVLGNATNKMAPQANLVFQSIMDSGGGLGGLPSNLQTLFSGQAFS 119
QY 121 AGARIHTNSWGAANVAGYTTDSRVVDYVRKNDMTILFAAGNEGPNQGTISAPGTAKNAI 180
DB 121 AGARIHTNSWGAAPVNGAYTTDSRVVDYVRKNDMTILFAAGNEGPNQGTISAPGTAKNAI 179
QY 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTKGRIPDVMAQPTIYLSARSSLAPDSGF 240
DB 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTKGRIPDVMAQPTIYLSARSSLAPDSGF 239
QY 241 WANDSKKAYMGGSMTPIVAGVNAQLREHFVNRGITPKPSILRAALTAGADVGLGY 300
DB 241 WANDSKKAYMGGSMTPIVAGVNAQLREHFVNRGITPKPSILRAALTAGADVGLGY 299
QY 301 PNGNQGWGRVTLDSKLVNAVYNESSALSTSOKATYFTATAGKPLKISLWSDAPASTTA 360
DB 301 PNGNQGWGRVTLDSKLVNAVYNESSALSTSOKATYFTATAGKPLKISLWSDAPASTTA 359
QY 361 SVTLVNDLDTVITAPNGTRVYGNDFSAFPDNNMDGRNNVENVFINSQSGTYTVEQAYN 420
DB 361 SVTLVNDLDTVITAPNGTRVYGNDFSAFPDNNMDGRNNVENVFINSQSGTYTVEQAYN 419
QY 421 VPVGPNFSLAIYN 434
DB 421 VPVGPNFSLAIYN 434

```

```

Db          420 VPVSPOTFSLAI VH 433

RESULT 10
ID          AAM89547
XX          AAM89547 standard; protein; 641 AA.
XX          AAM89547;
XX          AAM89547;
DT          12-APR-1999 (first entry)
XX          Bacillus JPI170 protease.
XX          Bacillus JPI170 protease.
XX          Protease; detergent; surfactant; leather processing; delectering;
XX          flavour.
XX          Bacillus sp.
XX          Key
XX          Location/Qualifiers
XX          Peptide
XX          1..33
XX          /note="signal peptide"
XX          Region
XX          34..208
XX          /note="prepro region"
XX          Protein
XX          209..641
XX          /note="mature protein"
XX          FT
XX          FT
XX          PN          WO9856927-A2.
XX          PD          17-DEC-1998.
XX          PF          09-JUN-1998; 98WO-US012005.
XX          PR          12-JUN-1997; 97US-00873479.
XX          PA          (NOVO ) NOVO NORDISK BIOTECH INC.
XX          PI          Sloma A, Christianson L;
XX          DR          WPI; 1999-080908/07.
XX          DR          N-PSDB; AAV82382.
XX          PT          Novel protease from Bacillus subtilis LC20 - useful in laundry and
XX          PT          dishwashing detergents and for leather processing.
XX          PS          Claim 7, Page 53-54; 77pp; English.
XX          CC          This is the amino acid sequence of a novel protease of Bacillus sp. JPI170
XX          CC          (NCIB 12513), as deduced from the nucleotide sequence of an isolated gene
XX          CC          (see AAV82382). The entire protein, including the signal peptide and
XX          CC          prepro region, has 77% identity to alkaline protease Y (see AAM89548)
XX          CC          from Bacillus. The invention provides vectors, recombinant host cells and
XX          CC          methods for the recombinant production of the protease. The protease is
XX          CC          used in laundry and dishwashing detergents, for institutional and
XX          CC          industrial cleaning, and for leather processing, as well as for
XX          CC          delectering and enhancing the degree of hydrolysis of protein
XX          CC          hydrolysates, for flavour development through hydrolysis of proteins.
XX          CC          It has enhanced stability towards oxidation under alkaline conditions.
XX          CC          e.g. towards bleaching agents of the peroxy type. The invention also
XX          CC          provides mutant cells in which the protease activity is diminished. Such
XX          CC          cells can be used for the production of heterologous recombinant proteins
XX          CC          SQ
XX          Sequence 641 AA:

Query Match          94.4%; Score 2118.5; DB 2; Length 641;
Best Local Similarity 93.3%; Pred. No. 35-146;
Matches 405; Conservative 18; Mismatches 10; Indels 1; Gaps 1

1 NDVARGIVKADVAQSSYGLVGGGQIVAAVDLTLDTRGDSMHEAFRGKITAYALGRITN 60
Db          209 NDVARGIVKADVAQSSYGLVGGGQIVAAVDLTLDTRGDSMHEAFRGKITAYALGRITN 268
61 NANDTNGHGVAVSGVLGNCATNKGAPOANLVFQSIWDSGGVGLGLPSNLTQLFSQAF 120

```

Db	269	NANDPNGHGTVAQSVLGN -ATNKGMA PQANLVFQSIIMDSGGIGGLPANIQLTFLSQAYS	327
Qy	121	AGAIHTNSKGAAYNGAYTTDSRVVDVYRKNDMTLLFAAGNEBPNGGTISAPGTAKAI	180
Db	328	AGARIHTNSWAPNGAAYTTDSRVVDVYRKNDMTLLFAAGNEBPGSGTISAPGTAKAI	387
Qy	181	TVGATENLRPFQSGYADANINNHVAQFSSRGPTKDGRIKPDVWAPGYILSARSLAPSSSF	240
Db	388	TVGATENLRPFQSGYADANINNHVAQFSSRGPTKDGRIKPDVWAPGYILSARSLAPSSSF	447
Qy	241	WANHDSKTYANGTGSMTAPVIAGNVAOLREHFVNKRGITTPRPSLLKALIGAADVGI	300
Db	448	WANHDSKTYANGTGSMTAPVIAGNVAOLREHFVNKRGITTPRPSLLKALIGAADVGI	507
Qy	301	PNGHOGKRVTLDSLSLVAYVNESSALSTSGATYFTATAGKPLKISLVMSDAPASTTA	360
Db	508	PNGHOGKRVTLDSLSLVAYVNESSALSTSGATYFTATAGKPLKISLVMSDAPASTTA	567
Qy	361	SVTLVNDLDELITAPNGTRVYVNDPSAPFCDNMWDGRNNVENVFINSPOSGTYTTEVQAYN	420
Db	568	SLTLVNDLDELITAPNGTRVYVNDPSAPFCDNMWDGRNNVENVFINSPOSGTYTTEVQAYN	627
Qy	421	VPVGPQNPSLAIYN 434	
Db	628	VPVSPQTSLAIYH 641	
RESULT 11			
ID	AAMS0090	AAMS0090 standard; protein: 434 AA.	
AC	AAMS0090;		
XX	12-AUG-2002	(first entry)	
DT			
XX			
DE	Bacillus sp KSM-KP43 alkaline protease protein variant.		
XX			
KW	Alkaline protease; detergent; laundry; bleaching; dishwasher; mutant;		
KM	mutcin.		
XX			
OS	Bacillus sp.		
XX			
XX	Synthetic.		
XX			
PH	Key	Location/Qualifiers	
FT	Misc-difference	46	/label= 'y,w,a,d,e,t,v,l,i,h,s,k,q,m,c
FT		/note= "as claimed in Claim 3"	
FT	Misc-difference	54	/label= 'y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c
FT		/note= "as claimed in Claim 3"	
FT	Misc-difference	57	/label= 'k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a
FT		/note= "as claimed in Claim 3"	
FT	Misc-difference	66	/label= 'e,d,s,g,a,t,l,m,c,v,g,i
FT		/note= "as claimed in Claim 3"	
FT	Misc-difference	84	/label= OTHER, R
FT		/note= "OTHER= deleted residue. Specifically described in Claim 1"	
FT	Misc-difference	101, 106	/label= 'k,s,e,f,v,r,y,l,i,t,m,c,y,d,e,h,p,a
FT		/note= "as claimed in Claim 3"	
FT	Misc-difference	104	/label= OTHER, P
FT		/note= "OTHER= deleted residue. Specifically described in Claim 1"	
FT	Misc-difference	107	/label= 'k,r,a,s
FT		/note= "as claimed in Claim 3"	
FT	Misc-difference	119	/label= 'y,f,a,n,e,t,v,h,s,k,e,m,g,d,p,r,c
FT			

[illegible]

CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency %
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease SD-521 from the
 CC *Bacillus* sp strain SD-521 (FERM BP-11162) described in the method of the
 CC invention

XX SQ Sequence 433 AA;

Query Match 89.6%; Score 2010.5; DB 5; Length 433;
 Best Local Similarity 88.5%; Pred. No. 14e-138;
 Matches 384; Conservative 26; Mismatches 23; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALVALGRIN 60
 DB 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALVALGRIN 60
 QY 61 NANDINGHGTHTVAGSVLNGATNKGMAPQANLVFQSIIMDSGGGLGSLPSNLTLPFQAFS 120
 DB 61 NANDINGHGTHTVAGSVLNGATNKGMAPQANLVFQSIIMDSGGGLGSLPSNLTLPFQAFS 119
 QY 121 AGARHTNSWGAAVNGAYTTDSRNVDYRKNDMTILFAAGNEGPNGGTISAPGTAKNAI 180
 DB 120 AGARHTNSWGAAPVNGAYTANSRQVDEYRNNDMTVLFAAGNEGPNGGTISAPGTAKNAI 179
 QY 181 TVGATENTLRFPSGSYADNINHYAOTSSRGPTKDGRIKPDVMAFGTYILSARSSLAPDSSE 240
 DB 180 TVGATENTLRFPSGSYADNINHYAOTSSRGPTKDGRIKPDVMAFGTYILSARSSLAPDSSE 239
 QY 241 MANHDSKYAYMGTSNATPIVAGNVAQLREHFVKRIGTIPKPSILKALILAGADVGLGY 300
 DB 240 MANHDSKYAYMGTSNATPIVAGNVAQLREHFVKRIGTIPKPSILKALILAGADVGLGY 299
 QY 301 PNGNQGWRVTLDKSLNVAAYNESSALSTSQKATYTFITATAGKPLKISLWSDAPASTTA 360
 DB 300 PNGNQGWRVTLDKSLNVAAYNESSALSTSQKATYTFITATAGKPLKISLWSDAPASTTA 359
 QY 361 SYTLVNDLDTVITAPNGTRYVGNDFAPFDNNMDGNNVENYFINSPOGGTYTIEVOAYN 420
 DB 360 SYTLVNDLDTVITAPNGTRYVGNDFAPFDNNMDGNNVENYFINSPOGGTYTIEVOAYN 419
 QY 421 VPVGPQNFSLAIYN 434
 DB 420 VPVGPQNFSLAIYN 433

RESULT 13

AA050082 ID AA050082 standard; protein; 433 AA.

XX AC AA050082;

DT 12-ATG-2002 (first entry)

DE *Bacillus* sp D6-(FERM P1592) alkaline protease protein fragment.

XX Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX OS *Bacillus* sp.

XX PN EP1209233-A2.

XX PD 29-MAY-2002.

XX PF 22-NOV-2001; 2001BP-00127851.

XX PR 22-NOV-2000; 2000JP-00355166.

XX PR 12-APR-2001; 2001JP-00114048.

XX PA (KAOS) KAO CORP.

XX PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N,
 PI Okuda M, Sasaki K;

XX WPI; 2002-437518/47.

XX PT New modified alkaline proteases useful in detergent compositions.

XX PS Claim 5; Page 13-15; 25pp; English.

XX This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency %
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease F-1 from *Bacillus*
 CC sp strain D6-(FERM-P1592) described in the method of the invention

XX SQ Sequence 433 AA;

Query Match 89.4%; Score 2006.5; DB 5; Length 433;
 Best Local Similarity 88.2%; Pred. No. 2.8e-138;
 Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALVALGRIN 60
 DB 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALVALGRIN 60
 QY 61 NANDINGHGTHTVAGSVLNGATNKGMAPQANLVFQSIIMDSGGGLGSLPSNLTLPFQAFS 120
 DB 61 NANDINGHGTHTVAGSVLNGATNKGMAPQANLVFQSIIMDSGGGLGSLPSNLTLPFQAFS 119
 QY 121 AGARHTNSWGAAVNGAYTTDSRNVDYRKNDMTILFAAGNEGPNGGTISAPGTAKNAI 180
 DB 120 AGARHTNSWGAAPVNGAYTANSRQVDEYRNNDMTVLFAAGNEGPNGGTISAPGTAKNAI 179
 QY 181 TVGATENTLRFPSGSYADNINHYAOTSSRGPTKDGRIKPDVMAFGTYILSARSSLAPDSSE 240
 DB 180 TVGATENTLRFPSGSYADNINHYAOTSSRGPTKDGRIKPDVMAFGTYILSARSSLAPDSSE 239
 QY 241 MANHDSKYAYMGTSNATPIVAGNVAQLREHFVKRIGTIPKPSILKALILAGADVGLGY 300
 DB 240 MANHDSKYAYMGTSNATPIVAGNVAQLREHFVKRIGTIPKPSILKALILAGADVGLGY 299
 QY 301 PNGNQGWRVTLDKSLNVAAYNESSALSTSQKATYTFITATAGKPLKISLWSDAPASTTA 360
 DB 300 PNGNQGWRVTLDKSLNVAAYNESSALSTSQKATYTFITATAGKPLKISLWSDAPASTTA 359
 QY 361 SYTLVNDLDTVITAPNGTRYVGNDFAPFDNNMDGNNVENYFINSPOGGTYTIEVOAYN 420
 DB 360 SYTLVNDLDTVITAPNGTRYVGNDFAPFDNNMDGNNVENYFINSPOGGTYTIEVOAYN 419
 QY 421 VPVGPQNFSLAIYN 434
 DB 420 VPVGPQNFSLAIYN 433

RESULT 14

AA089548 ID AA089548 standard; protein; 636 AA.

XX AC AA089548;

DT 12-APR-1999 (first entry)

DE *Bacillus* sp. alkaline protease Y.

XX Alkaline protease Y; detergent; surfactant; leather processing;
 XX debittering; flavour.

XX OS *Bacillus* sp.

XX PN WO9856927-A2.

XX PD 17-DEC-1998.

XX PF 09-JUN-1998; 98WO-US012005.

XX 12-JUN-1997; 97US-00873479.
 XX (NOVO) NOVO NORDISK BIOTECH INC.
 XX Sloma A, Christianson L;
 XX WPI, 1999-080908/07.
 PT Novel protease from *Bacillus subtilis* LC20 - useful in laundry and
 PT dishwashing detergents and for leather processing.
 XX Claim 3; Page 55-56; 77pp; English.
 CC This is the amino acid sequence of a *Bacillus* sp. alkaline protease Y
 CC that is said to have good alkali and surfactant resistance and improved
 CC detergency. It shows 77% identity to a newly isolated protease (see
 CC AAM89547) of *Bacillus* sp. JF170 (NCIB 12513). The invention provides
 CC vectors, recombinant host cells and methods for the recombinant
 CC production of such proteases. The protease are used in laundry and
 CC dishwashing detergents, for institutional and industrial cleaning, and
 CC for leather processing, as well as for debittering and enhancing the
 CC degree of hydrolysis of protein hydrolysates, for flavour development
 CC through hydrolysis of proteins, degradation of undesired peptides and in
 CC enzymatic synthesis of peptides. They have enhanced stability towards
 CC oxidation under alkaline conditions, e.g. towards bleaching agents of the
 CC peroxy type. The invention also provides mutant cells in which the
 CC protease activity is diminished. Such cells can be used for the
 CC production of heterologous recombinant proteins
 XX
 SQ Sequence 636 AA;
 Query Match 89.3%; Score 2004.5; D5 2; Length 636;
 Best Local Similarity 88.2%; Pred. No. 6.5e-138;
 Matches 383; Conservative 27; Mismatches 23; Indels 1; Gaps 1;
 QY 1 NDVARGIVKADVAQSSGLYGGQIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 60
 DB 204 NDVARGIVKADVAQNNNGLYGGQIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 263
 QY 61 NANDTNGHGTVAAGSVLGNATNKMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSGQAFS 120
 DB 264 NASDPNKGHTHVAGSVLGN-ALNKMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSGQAFN 322
 QY 121 AGARIHNSWGAANVGAYTTDSRRVDDYKAKDMTILFAAGNPGNGGTTISAPGTAKNAI 180
 DB 323 AGARIHNSWGAANVGAYTTDSRRVDDYKAKDMTILFAAGNPGNGGTTISAPGTAKNAI 382
 QY 181 TVGATENTLRSPFGSYADININHVAFSSRGPTKGRITKPDVMAFGTYILSARSSLAPDSSF 240
 DB 383 TVGATENTLRSPFGSYADININHVAFSSRGPTKGRITKPDVMAFGTYILSARSSLAPDSSF 442
 QY 241 WANHDSKYAVMGTSMTPTIVAGNVAQLREHFVNKRGITPKPSLKAALJAGADVGLGY 300
 DB 443 WANHDSKYAVMGTSMTPTIVAGNVAQLREHFVNKRGITPKPSLKAALJAGADVGLGY 502
 QY 301 PNGOQGWGRVTLDKSLVAAYVNESALSTOKATYTFETAAGKELISLWSPAPASTA 360
 DB 503 PNGOQGWGRVTLDKSLVAAYVNESALSTOKATYTFETAAGKELISLWSPAPASTA 562
 QY 361 SVTLVNDLDLVITAPNGTRVYGNDFSAFPDNNMGNNVENVINSPOGTYTIEVOAYN 420
 DB 563 SVTLVNDLDLVITAPNGTRVYGNDFSAFPDNNMGNNVENVINSPOGTYTIEVOAYN 622
 QY 421 VPVGPONFSLAIYN 434
 DB 623 VPVGPONFSLAIYN 636

AC AAM50083;
 XX 12-AUG-2002 (first entry)
 XX *Bacillus* sp. Y-(FERM BP-1029) alkaline protease protein fragment.
 DE Alkaline protease; detergent; laundry; bleaching; dishwasher.
 XX *Bacillus* sp.
 XX EPI209233-A2.
 XX 29-MAY-2002.
 XX 22-NOV-2001; 2001EP-00127851.
 XX 22-NOV-2000; 2000JP-00355166.
 XX 12-APR-2001; 2001JP-00114048.
 XX (KAO) KAO CORP.
 XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Arai H, Sumitomo N;
 PT Okuda M, Saeki K;
 DR WPI, 2002-437518/47.
 PT New modified alkaline proteases useful in detergent compositions.
 XX Claim 5; Page 15-16; 25pp; English.
 CC This invention describes novel *Bacillus* sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergency
 CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
 CC sequence represents a fragment of the alkaline protease Ya from *Bacillus*
 CC sp. strain Y-(FERM BP-1029) described in the method of the invention
 XX
 SQ Sequence 433 AA;
 Query Match 89.1%; Score 1999.5; D5 5; Length 433;
 Best Local Similarity 88.0%; Pred. No. 9e-138;
 Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;
 QY 1 NDVARGIVKADVAQSSGLYGGQIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 60
 DB 1 NDVARGIVKADVAQNNNGLYGGQIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 60
 QY 61 NANDTNGHGTVAAGSVLGNATNKMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSGQAFS 120
 DB 61 NASDPNKGHTHVAGSVLGN-ALNKMAPQANLVFQSIIMDSGGGLGGLPSNLQTLFSGQAFN 119
 QY 121 AGARIHNSWGAANVGAYTTDSRRVDDYKAKDMTILFAAGNPGNGGTTISAPGTAKNAI 180
 DB 120 AGARIHNSWGAANVGAYTTDSRRVDDYKAKDMTILFAAGNPGNGGTTISAPGTAKNAI 179
 QY 181 TVGATENTLRSPFGSYADININHVAFSSRGPTKGRITKPDVMAFGTYILSARSSLAPDSSF 240
 DB 180 TVGATENTLRSPFGSYADININHVAFSSRGPTKGRITKPDVMAFGTYILSARSSLAPDSSF 229
 QY 241 WANHDSKYAVMGTSMTPTIVAGNVAQLREHFVNKRGITPKPSLKAALJAGADVGLGY 300
 DB 240 WANHDSKYAVMGTSMTPTIVAGNVAQLREHFVNKRGITPKPSLKAALJAGADVGLGY 299
 QY 301 PNGOQGWGRVTLDKSLVAAYVNESALSTOKATYTFETAAGKELISLWSPAPASTA 360
 DB 300 PNGOQGWGRVTLDKSLVAAYVNESALSTOKATYTFETAAGKELISLWSPAPASTA 359
 QY 361 SVTLVNDLDLVITAPNGTRVYGNDFSAFPDNNMGNNVENVINSPOGTYTIEVOAYN 420
 DB 360 SVTLVNDLDLVITAPNGTRVYGNDFSAFPDNNMGNNVENVINSPOGTYTIEVOAYN 419
 QY 421 VPVGPONFSLAIYN 434

Sat Apr 3 17:19:58 2004

us-09-985-689a-2.rag

Page 11

DB 420 VPSGPRFSLAIVH 433

Search completed: March 31, 2004, 16:04:29
Job time : 49.2392 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: March 31, 2004, 15:59:39 ; Search time 14.0215 Seconds
(without alignments)
1597.947 Million cell updates/sec

Title: US-09-985-689A-2

Perfect score: 2245

Sequence: 1 NDVARGIVKADVAGSSVGLY.....EVQAVNVPVGPQNSLAIVN 434

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Listing first 45 summaries

Database: Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/CTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2237	99.6	639	US-09-509-814A-4	Sequence 4, Appli
2	2197	97.9	640	US-09-509-814A-8	Sequence 8, Appli
3	2181	97.6	640	US-09-509-814A-6	Sequence 6, Appli
4	2155	96.0	639	US-09-509-814A-1	Sequence 1, Appli
5	2155	96.0	640	US-09-509-814A-2	Sequence 2, Appli
6	2118.5	94.4	641	US-08-873-479-42	Sequence 42, Appli
7	1998.5	89.0	433	US-09-104-623A-4	Sequence 4, Appli
8	1998.5	89.0	433	US-09-019-532-4	Sequence 4, Appli
9	1998.5	89.0	433	US-09-338-746-4	Sequence 4, Appli
10	1998.5	89.0	635	US-08-873-479-43	Sequence 43, Appli
11	1579.5	70.4	345	US-09-512-251A-10	Sequence 10, Appli
12	1579.5	70.4	345	US-09-515-150A-10	Sequence 10, Appli
13	1579.5	70.4	345	US-09-196-281-13	Sequence 13, Appli
14	1579.5	70.4	345	US-08-894-818B-1	Sequence 1, Appli
15	1579.5	70.4	345	US-08-894-818B-1	Sequence 1, Appli
16	1579.5	70.4	345	US-09-445-472-12	Sequence 12, Appli
17	1579.5	70.4	345	US-09-445-472-1	Sequence 1, Appli
18	1579.5	70.4	345	US-08-894-818B-3	Sequence 3, Appli
19	1579.5	70.4	345	US-09-445-472-4	Sequence 4, Appli
20	1579.5	70.4	345	US-08-894-818B-5	Sequence 5, Appli
21	1579.5	70.4	345	US-09-445-472-16	Sequence 16, Appli
22	1579.5	70.4	345	US-08-894-818B-5	Sequence 5, Appli
23	1579.5	70.4	345	US-09-000-016-7	Sequence 7, Appli
24	1579.5	70.4	345	US-09-514-340-7	Sequence 7, Appli
25	1579.5	70.4	345	US-09-000-016-4	Sequence 4, Appli
26	1579.5	70.4	345	US-09-514-340-4	Sequence 4, Appli
27	1579.5	70.4	345	US-09-000-016-2	Sequence 2, Appli
28	1579.5	70.4	345	US-09-514-340-2	Sequence 2, Appli

28	304.5	13.6	903	1	US-08-750-532-1	Sequence 1, Appli
29	304.5	13.6	1398	1	US-08-750-532-9	Sequence 9, Appli
30	304.5	13.6	1398	3	US-08-894-818B-8	Sequence 8, Appli
31	304.5	13.6	1398	4	US-09-445-472-6	Sequence 6, Appli
32	296.5	13.2	237	1	US-08-750-532-18	Sequence 18, Appli
33	285.5	12.7	418	2	US-08-873-479-44	Sequence 44, Appli
34	273	12.2	418	4	US-09-966-921A-2	Sequence 2, Appli
35	260.5	11.6	397	4	US-08-328-352-7533	Sequence 7533, Ap
36	259.5	11.6	275	1	US-08-431-387-1	Sequence 1, Appli
37	256	11.4	280	1	US-08-434-255-8	Sequence 8, Appli
38	256	11.4	280	1	US-08-459-967-8	Sequence 8, Appli
39	256	11.4	280	1	US-08-460-327-8	Sequence 8, Appli
40	256	11.4	280	1	US-08-459-871-8	Sequence 8, Appli
41	256	11.4	280	3	US-09-024-532-2	Sequence 2, Appli
42	256	11.4	280	4	US-09-104-623A-2	Sequence 2, Appli
43	256	11.4	280	4	US-09-019-532-2	Sequence 2, Appli
44	256	11.4	280	4	US-09-417-359A-2	Sequence 2, Appli
45	256	11.4	280	4	US-09-705-185-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1	US-09-509-814A-4	
Sequence 4, Application	US/09509814A	
Patent No. 6376227		
GENERAL INFORMATION:		
APPLICANT: TAKAIIWA, MIKIO		
APPLICANT: OKUDA, MITSUYOSHI		
APPLICANT: SAKI, KATSUHIKA		
APPLICANT: KIMOTO, HIROMI		
APPLICANT: HITOMI, JUN		
APPLICANT: KAGEYAMA, YASUSHI		
APPLICANT: SHIKATA, SHITSUM		
APPLICANT: NOMURA, MASAFUMI		
TITLE OF INVENTION: ALKALINE PROTEASE		
FILE REFERENCE: 0327-0832-OPCT		
CURRENT APPLICATION NUMBER: US/09/509,814A		
CURRENT FILING DATE: 2000-04-06		
PRIOR APPLICATION NUMBER: PCT/JP98/04528		
PRIOR FILING DATE: 1998-10-07		
PRIOR APPLICATION NUMBER: JP 9-274570		
PRIOR FILING DATE: 1997-06-08		
NUMBER OF SEQ ID NOS: 24		
SOFTWARE: PatentIn version 3.0		
SEQ ID NO 4		
LENGTH: 639		
TYPE: PRT		
ORGANISM: Bacillus sp.		
US-09-509-814A-4		
Query Match	99.6%	Score 2237; DB 4; Length 639;
Best Local Similarity	99.8%	Pred. No. 7.4e-175;
Matches 433; Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
1 NDVARGIVKADVAGSSVGLYGGQIIVAVADTGIDTGRNDSNMEAFRGKITALYALGRIN 60		
206 NDVARGIVKADVAGSSVGLYGGQIIVAVADTGIDTGRNDSNMEAFRGKITALYALGRIN 265		
61 NADNTHGHVHVGSTLNGATKMGAPANLVFQSSIMSSGGLPENTLTFEQARS 120		
266 NADNTHGHVHVGSTLNGATKMGAPANLVFQSSIMSSGGLPENTLTFEQARS 325		
121 AGARHTNSGAAVNGAYTTDSRNVDDYVRKDMTILFAAGNEGPGGTISAPGTAKNAI 180		
326 AGARHTNSGAAVNGAYTTDSRNVDDYVRKDMTILFAAGNERPGGTISAPGTAKNAI 385		
181 TVGATNTLRPSFSSVADNINHTVAQFSSRGPTDGRKIPVMAPGYTLTARSISAPDSGF 240		
386 TVGATNTLRPSFSSVADNINHTVAQFSSRGPTDGRKIPVMAPGYTLTARSISAPDSGF 445		
241 WANHSKRAYAGMGTSNATPIVAGNVAGLREHFVKRGTIPKPSLLKALIIAGAADVGLGY 300		

```
Db 446 WANDSKAYWGTSMATPIVAGNVAQLREHFVKNRITPKPSLTKALITAGADVLGY 505
Qy 301 PNGQGWGRVTLDSLVNAVYNESSALSTQKATYTTATAGKPLKISLVMSDAPASTTA 360
Db 506 PNGQGWGRVTLDSLVNAVYNESSALSTQKATYTTATAGKPLKISLVMSDAPASTTA 565
Qy 361 SVTLVNDLDVITAPNSTRVYGNDFSAFPDNNMGRNNVENVFINSQSGTYTIEVOAYN 420
Db 566 SVTLVNDLDVITAPNSTRVYGNDFSAFPDNNMGRNNVENVFINSQSGTYTIEVOAYN 625
Qy 421 VPVGPQNFSLAIYN 434
Db 626 VPVGPQNFSLAIYN 639

RESULT 2
US-09-509-814A-8
; Sequence 8, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-8

Query Match 97.9%; Score 2197; DB 4; Length 640;
Best Local Similarity 96.8%; Pred. No. 1.4e-171;
Matches 420; Conservative 13; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NDVARGIVKADVAQSSYGLVGQGIIVAVDTGLDTSRNDSSMHEAFRGKITLALYALGRTN 60
Db 207 NDVARGIVKADVAQSSYGLVGQGIIVAVDTGLDTSRNDSSMHEAFRGKITLALYALGRTN 266
Qy 61 NANTDNGHGHVAGSVLGNATNKMAPQANLVFQSIMDSGGGLGJLPSNLQTLFSGQAFS 120
Db 267 NANTDNGHGHVAGSVLGNATNKMAPQANLVFQSIMDSGGGLGJLPSNLQTLFSGQAFS 326
Qy 121 AGARIHNSWGAANVNGAYTTDSRVVDYRKNDMTILFAAGNBGPNGGTISAGTAKNAI 180
Db 327 AGARIHNSWGAANVNGAYTTDSRVVDYRKNDMTILFAAGNBGPNGGTISAGTAKNAI 386
Qy 181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSF 240
Db 387 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSF 446
Qy 241 WANDSKAYWGTSMATPIVAGNVAQLREHFVKNRITPKPSLTKALITAGADVLGY 300
Db 447 WANDSKAYWGTSMATPIVAGNVAQLREHFVKNRITPKPSLTKALITAGADVLGY 506
Qy 301 PNGQGWGRVTLDSLVNAVYNESSALSTQKATYTTATAGKPLKISLVMSDAPASTTA 360
Db 507 PNGQGWGRVTLDSLVNAVYNESSALSTQKATYTTATAGKPLKISLVMSDAPASTTA 566
```

```
Qy 361 SVTLVNDLDVITAPNSTRVYGNDFSAFPDNNMGRNNVENVFINSQSGTYTIEVOAYN 420
Db 567 SVTLVNDLDVITAPNSTRVYGNDFSAFPDNNMGRNNVENVFINSQSGTYTIEVOAYN 626
Qy 421 VPVGPQNFSLAIYN 434
Db 627 VPVGPQNFSLAIYN 640

RESULT 3
US-09-509-814A-6
; Sequence 6, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-6

Query Match 97.6%; Score 2191; DB 4; Length 640;
Best Local Similarity 96.5%; Pred. No. 4.3e-171;
Matches 419; Conservative 13; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NDVARGIVKADVAQSSYGLVGQGIIVAVDTGLDTSRNDSSMHEAFRGKITLALYALGRTN 60
Db 207 NDVARGIVKADVAQSSYGLVGQGIIVAVDTGLDTSRNDSSMHEAFRGKITLALYALGRTN 266
Qy 61 NANTDNGHGHVAGSVLGNATNKMAPQANLVFQSIMDSGGGLGJLPSNLQTLFSGQAFS 120
Db 267 NANTDNGHGHVAGSVLGNATNKMAPQANLVFQSIMDSGGGLGJLPSNLQTLFSGQAFS 326
Qy 121 AGARIHNSWGAANVNGAYTTDSRVVDYRKNDMTILFAAGNBGPNGGTISAGTAKNAI 180
Db 327 AGARIHNSWGAANVNGAYTTDSRVVDYRKNDMTILFAAGNBGPNGGTISAGTAKNAI 386
Qy 181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSF 240
Db 387 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSF 446
Qy 241 WANDSKAYWGTSMATPIVAGNVAQLREHFVKNRITPKPSLTKALITAGADVLGY 300
Db 447 WANDSKAYWGTSMATPIVAGNVAQLREHFVKNRITPKPSLTKALITAGADVLGY 506
Qy 301 PNGQGWGRVTLDSLVNAVYNESSALSTQKATYTTATAGKPLKISLVMSDAPASTTA 360
Db 507 PNGQGWGRVTLDSLVNAVYNESSALSTQKATYTTATAGKPLKISLVMSDAPASTTA 566
Qy 361 SVTLVNDLDVITAPNSTRVYGNDFSAFPDNNMGRNNVENVFINSQSGTYTIEVOAYN 420
Db 567 SVTLVNDLDVITAPNSTRVYGNDFSAFPDNNMGRNNVENVFINSQSGTYTIEVOAYN 626
Qy 421 VPVGPQNFSLAIYN 434
Db 627 VPVGPQNFSLAIYN 640
```

```
RESULT 4
US-09-509-814A-1
Sequence 1, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKKI, KATSUSHISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 639
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (23)..(23)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (29)..(29)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (32)..(32)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (46)..(46)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (53)..(53)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (70)..(70)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (74)..(74)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (89)..(89)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (102)..(102)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (105)..(105)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (128)..(128)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (130)..(130)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (594)..(594)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
```

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1

Query Match 96.0%; Score 2155; DB 4; Length 639;
Best Local Similarity 96.3%; Pred. No. 3.8e-168;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAGSSYGLYGQGIYAVADTG:DTGRNDSMHEARFKITLALYALGRIN 60
DB 206 NDVARGIVKADVAGSSYGLYGQGIYAVADTG:DTGRNDSMHEARFKITLALYALGRIN 265
QY 61 NANDTNGHGHVAGSVLGNATNKMAPQANLVFQSIMSSGGLGGLPSNLQTLFQOAFS 120
DB 266 NANDTNGHGHVAGSVLGNATNKMAPQANLVFQSIMSSGGLGGLPSNLQTLFQOAFS 325
QY 121 AGARHTNSWGAANVGAYTTDSNRVDYRKNDMTLLFAAGNGEPNGGTTISAGTAKNAI 180
DB 326 AGARHTNSWGAANVGAYTTDSNRVDYRKNDMTLLFAAGNGEPNGGTTISAGTAKNAI 385
QY 181 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKEDVMAEGTYILSARSSLAPDSGF 240
DB 386 TVGATENLRPSFGSYADNINHVAQFSSRGPTKDGRIKEDVMAEGTYILSARSSLAPDSGF 445
QY 241 WANHDSKYAYVGGTSMATPIYAGVAVQLREHFKVNGRITPKPBLLAALAGADVGLGY 300
DB 446 WANHDSKYAYVGGTSMATPIYAGVAVQLREHFKVNGRITPKPBLLAALAGADVGLGY 505
QY 301 PNGNGGGRVTLDKSLNVAAYNESSALSTQKATYFTATAGKPLKISLVMSAPASTTA 360
DB 506 PNGNGGGRVTLDKSLNVAAYNESSALSTQKATYFTATAGKPLKISLVMSAPASTTA 565
QY 361 SVTLVNLDLVITTPNGTRYVGNDFSAFPNNMDGRNNVENVTINSPOSGTYTIEVOAYN 420
DB 566 SVTLVNLDLVITTPNGTRYVGNDFSAFPNNMDGRNNVENVTINSPOSGTYTIEVOAYN 625
QY 421 VPVGPONFSLAIYN 434
DB 626 VPVGPONFSLAIYN 639

RESULT 5
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUMI
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509,814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2

LENGTH: 640
TYPE: prt
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (124)..(124)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid

NAME/KEY: misc.feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc.feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-2

Query Match 96.0%; Score 2155; DB 4; Length 640;
Best Local Similarity 96.3%; Pred. No. 3,8e-168;
Matches 418; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 1 NDVARGIVADVAQSSYGLYGGQQLVAVADTGLDGRNDSNHEAFRGKITLVLALGRN 60
DB 207 NDVARGIVADVAQSSYGLYGGQQLVAVADTGLDGRNDSNHEAFRGKITLVLALGRN 266
QY 61 NANDINGHGHVAGSVLGNATNKGAPANLVFOSISGGLGSLPSNLQTLFSGAAS 120
DB 267 NANDINGHGHVAGSVLGNATNKGAPANLVFOSISGGLGSLPSNLQTLFSGAAS 326

QY 121 AGARITNSWGAANVAYTTDSRNVDYVRKNDMTLFLAAGNEGPNGTISAPGAKAI 180
DB 327 AGARITNSWGAANVAYTTDSRNVDYVRKNDMTLFLAAGNEGPNGTISAPGAKAI 386
QY 181 TVGATENLRPSFGSVADNINNVAAQFSSRGPTDGRKPDVAPGYIISARSSLPDSSF 240
DB 387 TVGATENLRPSFGSVADNINNVAAQFSSRGPTDGRKPDVAPGYIISARSSLPDSSF 446
QY 241 WANHSKRYVMGTSVATPIVAGNVAKLREHFKVRGKITPKSLKALILAGADVGLCY 300
DB 447 WANHSKRYVMGTSVATPIVAGNVAKLREHFKVRGKITPKSLKALILAGADVGLCY 506
QY 301 PGNQMGRTVLDKSLNVAVYNESSALSTOKATYFTATAGKPKLSLWSDAPASTTA 360
DB 507 PGNQMGRTVLDKSLNVAVYNESSALSTOKATYFTATAGKPKLSLWSDAPASTTA 566
QY 361 SVTLVNDLIVITAPRGTYVGNDFAPDNMMDGNNTNENFINSPOGCTTIEVOAYN 420
DB 567 SVTLVNDLIVITAPRGTYVGNDFAPDNMMDGNNTNENFINSPOGCTTIEVOAYN 626
QY 421 VPVGPONFSLAIVN 434
DB 627 VPVGPONFSLAIVN 640

RESULT 6
US-08-873-479-42
Sequence 42; Application US/08873479
Patent No. 5891701

GENERAL INFORMATION:

APPLICANT: Sloana, Alan
TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
TITLE OF INVENTION: Having Protease Activity
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5891701 of No. 5891701 of America
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/873,479
FILING DATE: 12-JUN-1997
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Agitis, Cheryl H
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 5251,000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:

INFORMATION FOR SEQ ID NO: 42:

SEQUENCE CHARACTERISTICS:
LENGTH: 641 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-873-479-42

Query Match 94.4%; Score 2118.5; DB 2; Length 641;
Best Local Similarity 93.3%; Pred. No. 3,7e-165;
Matches 405; Conservative 18; Mismatches 10; Indels 1; Gaps 1;

```

;       TOPOLOGY: linear
;       MOLECULE TYPE: Protein
;       ORIGINAL SOURCE:
;       STRAIN: Bacillus sp. Y
US-09-104-623A-4

Query Match      89.0%; Score 1998.5; DB 4; Length 433;
Best Local Similarity 88.0%; Pred. No. 1.4e-15;
Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

QY      1 NDVARGIVKADVAAQSYGLVGQGGIVAVADPTGDTGRNDSMEBAFPGKITALVALGRTN    60
DB      1 NDVARGIVKADVAAQNNTGLYGQSGLVAVDITGLDGTGRNDSMEBAFFGKITALLALGRTN    60
QY      : :::::::::::::::::::::
DB      61 NASDPNGHGTHVASVIGNGATNKGNAPQANLVFOSIMDSGGIGLIPSNLTQLFSQAIFS    119
QY      : :::::::::::::::::::::
DB      121 AGADITHNSMGAANAAGAYTTDSRVADDYVRKNMDTILFAAGNEGPNCGTISAPGTAKNAI    180
QY      120 AGAIIHNSMGAPNGAYTANSRVDYEVRRNDMTVLFAAGNEPNSGTSISAPTAKNAI    179
DB      181 TVGATEMLPSFGSVADNNINHHVAQFPSSRGPTKGRIKPDVMAPEGTYLISARSIAPDSF    240
DB      180 TVGATEVPRSPFGSIALNPNIHIAQFPSSRGATRDRIRIKPDVTAPETFILSARSIAPDSF    239
QY      241 WANIDSKYAAYVGGTSNMATPIYAGVAOLREHFVNKRGIPTPPSLTKLAALAGADVGLGY    300
DB      240 WANNYSKAAYVGGTSNMATPIYAGVAOLREHFINKRGITTPPSLIKANLLAGADVGLGY    299
QY      301 PNGSGWGVTLDKSLNVAAYNESSALSTOKATYTFTATGKPLKLSIWSDAPASTTA    360
DB      300 PSGDGWGVRTLDKSLNVAAYVENTALTATGKATYFQAOQKGPDLKLSLWTTDAPGSTTA    359
QY      361 SVTLVNDLDVLTLPNGNRVYGNPFSAPEFDNMNGRNVEVFINSPOSCTYTLAEVQAYN    420
DB      360 SYTLVNDLDVLTLPNGOKRYGDNFSTFYDNNMDGRNVEVFINFAPSGTYITIEVQAYN    419
QY      421 VPVGPQNFSLAIYN 434
DB      420 VPSGPFQRFSIAIVH 433

RESULT 8
US-09-019-532-4
; Sequence 4, Application US/09019532B
; Patent No. 6416756
; GENERAL INFORMATION:
; APPLICANT: Prent, Annette
; TITLE OF INVENTION: A Modified Enzyme for Skin Care
; FILE REFERENCE: 4922.204-US
; CURRENT APPLICATION NUMBER: US/09/019,532B
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: 0036/97
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 0754/97
; EARLIER FILING DATE: 1997-06-25
; EARLIER APPLICATION NUMBER: 60/051,381
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: PCT/DK98/00015
; EARLIER FILING DATE: 1998-01-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-019-532-4

Query Match      89.0%; Score 1998.5; DB 4; Length 433;
Best Local Similarity 88.0%; Pred. No. 1.4e-15;
Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

```

```

QY 1 NDVARGIVKADVAGSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRIN 60
DB 1 NDVARGIVKADVAGSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRIN 60
QY 61 NADDTGHTGHTVAGSVTLGNATKGMAPQANLVFQSIIMDSGGGLGSLPSNLQTLFSGQAFS 120
DB 61 NADDPHGHGHTVAGSVTLGN-ALNKGMAPQANLVFQSIIMDSGGGLGSLPSNLQTLFSGQAFS 119
QY 121 AGARIHNSWGAANVNGAYTTDSRNVDYVRKNDMTLFAAGNEGPNNGTISAPGTAKNAI 180
DB 120 AGARIHNSWGAANVNGAYTTANSRQVDEYVANNMTVLFAAGNEGPNNGTISAPGTAKNAI 179
QY 181 TVGATENLRPSFGSYADNINNHVAQFSSRGPTKDGRIKPDVMAFGTYILSARSSILAPDSSF 240
DB 180 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSILAPDSSF 239
QY 241 MANHDSKYAWMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLTKALIAAGADVGLGY 300
DB 240 MANYSKYAWMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLTKALIAAGADVGLGY 299
QY 301 PNGNQGWRVTLDKSLNVAAYVNESSALSTSQKATYFTATAGKPLKISLVMSDAPASTTA 360
DB 300 PNGDQGWGRVTLDKSLNVAAYVNEATATATGQKATYSFQAQKPLKISLVMTDAPASTTA 359
QY 361 SVTLVNDLDLVITAPNGTRYGNDFSAPEPNNDGRRNVNNTVINSQSGTYTIEVOAYN 420
DB 360 SVTLVNDLDLVITAPNGQKTVGNDFSYPDNNDGRRNVNNTVINSQSGTYTIEVOAYN 419
QY 421 VPVGPNFSLAIYN 434
DB 420 VPSGPQRFSLAIYH 433

```

RESULT 9

```

US-09-338-746-4
; Sequence 4, Application US/09338746
; Patent No. 6635526
; GENERAL INFORMATION:
; APPLICANT: Deussen, Heinz-Cosel
; APPLICANT: Olsen, Arne A.
; APPLICANT: Fatum, Tine M.
; APPLICANT: Roggen, Erwin L.
; TITLE OF INVENTION: A Polypeptide-Polymer Conjugate
; FILE REFERENCE: 5619,200-US
; CURRENT APPLICATION NUMBER: US/09/338,746
; EARLIER FILING DATE: 1999-06-23
; EARLIER APPLICATION NUMBER: PA 1998 00809
; EARLIER FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 60/091,461
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-338-746-4

```

```

Query Match 89.0%; Score 1998.5; DB 4; Length 433;
Best Local Similarity 88.0%; Pred. No. 1,4e-155;
Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

```

```

QY 1 NDVARGIVKADVAGSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRIN 60
DB 1 NDVARGIVKADVAGSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRIN 60
QY 61 NADDTGHTGHTVAGSVTLGNATKGMAPQANLVFQSIIMDSGGGLGSLPSNLQTLFSGQAFS 120
DB 61 NADDPHGHGHTVAGSVTLGN-ALNKGMAPQANLVFQSIIMDSGGGLGSLPSNLQTLFSGQAFS 119
QY 121 AGARIHNSWGAANVNGAYTTDSRNVDYVRKNDMTLFAAGNEGPNNGTISAPGTAKNAI 180
DB 120 AGARIHNSWGAANVNGAYTTANSRQVDEYVANNMTVLFAAGNEGPNNGTISAPGTAKNAI 179

```

```

DB 120 AGARIHNSWGAANVNGAYTTANSRQVDEYVANNMTVLFAAGNEGPNNGTISAPGTAKNAI 179
QY 181 TVGATENLRPSFGSYADNINNHVAQFSSRGPTKDGRIKPDVMAFGTYILSARSSILAPDSSF 240
DB 180 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSILAPDSSF 239
QY 241 MANHDSKYAWMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLTKALIAAGADVGLGY 300
DB 240 MANYSKYAWMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLTKALIAAGADVGLGY 299
QY 301 PNGNQGWRVTLDKSLNVAAYVNESSALSTSQKATYFTATAGKPLKISLVMSDAPASTTA 360
DB 300 PNGDQGWGRVTLDKSLNVAAYVNEATATATGQKATYSFQAQKPLKISLVMTDAPASTTA 359
QY 361 SVTLVNDLDLVITAPNGTRYGNDFSAPEPNNDGRRNVNNTVINSQSGTYTIEVOAYN 420
DB 360 SVTLVNDLDLVITAPNGQKTVGNDFSYPDNNDGRRNVNNTVINSQSGTYTIEVOAYN 419
QY 421 VPVGPNFSLAIYN 434
DB 420 VPSGPQRFSLAIYH 433

```

RESULT 10

```

US-08-873-479-43
; Sequence 43, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lytle, Christianson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESSES:
; ADDRESS: No. 58917010 No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,479
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agilis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251,000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-873-479-43

```

Query Match

```

Query Match 89.0%; Score 1998.5; DB 2; Length 635;
Best Local Similarity 88.0%; Pred. No. 2,4e-155;
Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

```

```

QY 1 NDVARGIVKADVAGSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRIN 60
DB 203 NDVARGIVKADVAGSSYGLYGQGIIVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRIN 262

```

```

QY 61 NANTNGHGHVAGSVLNGATNKMAPQANLVFQSIMDSGGGLPSNLQTLFSQAFS 120
Db 263 NADDPNGHGHVAGSVLGN-ALNKMAPQANLVFQSIMDSGGGLPSNLQTLFSQAFS 321
QY 121 AGARIHNSWGAANVNGAYTTDSRVNDYVRKNDMTILFAAGNPGNGTISAPGTAKNAI 180
Db 322 AGARIHNSWGAAPNGAYTANSRQVDEYVNNDMTILFAAGNPGNSGTISAPGTAKNAI 381
QY 181 TVGATEMLRPSFGSYAANINHVAFSSRGPTKGRIPKVMAFGTILSARSLAPDSGF 240
Db 382 TVGATEMLRPSFGSIADNPNHIAQFSSRGPTKGRIPKVMAFGTILSARSLAPDSGF 441
QY 241 WANHSKXYAMGTSMAFPIVAGNVAQLREHFVNKRGITPKPSILKALIAAGADVGLGY 300
Db 442 WANHSKXYAMGTSMAFPIVAGNVAQLREHFVNKRGITPKPSILKALIAAGADVGLGY 501
QY 301 PNGNGGWRVTLDKSLNVAAYVNSSALSTQKATYFTATAGPKIKISLVMSAPASTA 360
Db 502 PNGNGGWRVTLDKSLNVAAYVNEATATAGQATYSFQAGPKIKISLVMTDAPASTA 561
QY 361 SVTLVNDLVLITAPNGTRYVNDPSAPDNNDGNNVENVINSPOSQTYTIEVOAYN 420
Db 562 SVTLVNDLVLITAPNGQXVNDPSYPMNDGNNVENVINSPOSQTYTIEVOAYN 621
QY 421 VPVGPQNFSLATYN 434
Db 622 VPVGPQNFSLATYN 635

```

RESULT 11

```

US-09-512-251A-10
; Sequence 10, Application US/09512251A
; Patent No. 655355
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikelsen, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5349.204-US
; CURRENT APPLICATION NUMBER: US/09/512.251A
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
; US-09-512-251A-10

```

```

Query Match 70.4%; Score 1579.5; DB 4; Length 345;
Best Local Similarity 95.6%; Pred. No. 2e-121;
Matches 304; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

```

```

QY 1 NDVARGIVKADVAASSVGLYGGQGIYVADTGLDTGRNDSMEAFRGKITLALYALGRIN 60
Db 29 NDVARGIVKADVAAQNNGFLYGGQGIYVADTGLDTGRNDSMEAFRGKITLALYALGRIN 88
QY 61 NANTNGHGHVAGSVLNGATNKMAPQANLVFQSIMDSGGGLPSNLQTLFSQAFS 120
Db 89 NADDPNGHGHVAGSVLGN-ALNKMAPQANLVFQSIMDSGGGLPSNLQTLFSQAFS 321
QY 121 AGARIHNSWGAANVNGAYTTDSRVNDYVRKNDMTILFAAGNPGNGTISAPGTAKNAI 180
Db 148 AGARIHNSWGAAPNGAYTANSRQVDEYVNNDMTILFAAGNPGNSGTISAPGTAKNAI 381
QY 181 TVGATEMLRPSFGSYAANINHVAFSSRGPTKGRIPKVMAFGTILSARSLAPDSGF 240
Db 208 TVGATEMLRPSFGSYADNPNHIAQFSSRGPTKGRIPKVMAFGTILSARSLAPDSGF 441
QY 241 WANHSKXYAMGTSMAFPIVAGNVAQLREHFVNKRGITPKPSILKALIAAGADVGLGY 300
Db 268 WANHSKXYAMGTSMAFPIVAGNVAQLREHFVNKRGITPKPSILKALIAAGADVGLGY 501

```

```

QY 301 PNGNGGWRVTLDKSLNV 318
Db 328 PNGNGGWRVTLDKSLNV 345

```

RESULT 12

```

US-09-515-150A-10
; Sequence 10, Application US/09515150A
; Patent No. 655838
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikelsen, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5348.204-US
; CURRENT APPLICATION NUMBER: US/09/515.150A
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
; US-09-515-150A-10

```

```

Query Match 70.4%; Score 1579.5; DB 4; Length 345;
Best Local Similarity 95.6%; Pred. No. 2e-121;
Matches 304; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

```

```

QY 1 NDVARGIVKADVAASSVGLYGGQGIYVADTGLDTGRNDSMEAFRGKITLALYALGRIN 60
Db 29 NDVARGIVKADVAAQNNGFLYGGQGIYVADTGLDTGRNDSMEAFRGKITLALYALGRIN 88
QY 61 NANTNGHGHVAGSVLNGATNKMAPQANLVFQSIMDSGGGLPSNLQTLFSQAFS 120
Db 89 NADDPNGHGHVAGSVLGN-ALNKMAPQANLVFQSIMDSGGGLPSNLQTLFSQAFS 321
QY 121 AGARIHNSWGAANVNGAYTTDSRVNDYVRKNDMTILFAAGNPGNGTISAPGTAKNAI 180
Db 148 AGARIHNSWGAAPNGAYTANSRQVDEYVNNDMTILFAAGNPGNSGTISAPGTAKNAI 381
QY 181 TVGATEMLRPSFGSYAANINHVAFSSRGPTKGRIPKVMAFGTILSARSLAPDSGF 240
Db 208 TVGATEMLRPSFGSYADNPNHIAQFSSRGPTKGRIPKVMAFGTILSARSLAPDSGF 441
QY 241 WANHSKXYAMGTSMAFPIVAGNVAQLREHFVNKRGITPKPSILKALIAAGADVGLGY 300
Db 268 WANHSKXYAMGTSMAFPIVAGNVAQLREHFVNKRGITPKPSILKALIAAGADVGLGY 501
QY 301 PNGNGGWRVTLDKSLNV 318
Db 328 PNGNGGWRVTLDKSLNV 345

```

RESULT 13

```

US-09-196-281-13
; Sequence 13, Application US/09196281A
; Patent No. 660548
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter K.
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikelsen, Frank
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5435.200-US
; CURRENT APPLICATION NUMBER: US/09/196.281A
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: 1332/97
; EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows version 3.0
; SEQ ID NO 13

```

```

QY 301 PNGNGGWRVTLDKSLNV 318
Db 328 PNGNGGWRVTLDKSLNV 345

```


Sat Apr 3 17:19:59 2004

us-09-985-689a-2.rai

Page 9

```

/ LENGTH: 345
/ TYPE: PRT
/ ORGANISM: Bacillus
US-09-196-281-13

Query Match
Best Local Similarity 70.4%; Score 1579.5; DB 4; Length 345;
Matches 304; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

QY 1 NDVAGIYKADYAGSSYGLYGGQIYVAVDTGLDGRNDSNHEAFRGKITALYALGRTN 60
DB 29 NDVAGIYKADYAGSSYGLYGGQIYVAVDTGLDGRNDSNHEAFRGKITALYALGRTN 88
QY 61 NANDNGHGTAVAGSTLGNKGAPOANLVFOSIMDSGGLGGLPSNLOTFSQAFS 120
DB 89 NANDNGHGTAVAGSTLGN-ATNKGAPOANLVFOSIMDSGGLGGLPSNLOTFSQAFS 147
QY 121 AGARHTNSWGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPNGGTISAPGTAKXAI 180
DB 148 AGARHTNSWGAAPVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPNGGTISAPGTAKXAI 207
QY 181 TYGATENLRPSFGSYADNINNHVAFSSRGPTDGRIKPDVMAFGTYIISASSLAPDSF 240
DB 208 TYGATENLRPSFGSYADNINNHVAFSSRGPTDGRIKPDVMAFGTYIISASSLAPDSF 267
QY 241 WANHDSKYAMGTSMATPIVAGNVAQLREHFVNKRGITPKPSLIKALLIAGADVGIGY 300
DB 268 WANHDSKYAMGTSMATPIVAGNVAQLREHFVNKRGITPKPSLIKALLIAGADVGIGY 327
QY 301 PNGNQGMRVTLDKSLNV 318
DB 328 PNGNQGMRVTLDKSLNV 345

RESULT 14
US-08-894-818B-1
/ Sequence 1, Application US/08894818B
/ Patent No. 6261822
/ GENERAL INFORMATION:
/ APPLICANT: TAKAKURA, Hikaru
/ APPLICANT: MORISHITA, Mio
/ APPLICANT: YAMAMOTO, Katsuhiko
/ APPLICANT: MITTA, Masanori
/ APPLICANT: ASADA, Kiyozo
/ APPLICANT: TSUNASAWA, Susumu
/ APPLICANT: KATO, Ikunoshin
/ TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
/ NUMBER OF SEQUENCES: 42
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Broadway and Neimark
/ STREET: 419 Seventh Street N.W., Ste. 300
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: United States of America
/ ZIP: 20004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/894,818B
/ FILING DATE: 20-MAY-1998
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/JP96/03253
/ FILING DATE: 07-NOV-1996
/ PRIOR APPLICATION DATA: JP 333285/1995
/ APPLICATION NUMBER: 12-DEC-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Browdy, Roger L.
/ REGISTRATION NUMBER: 25,618
```

```

/ REFERENCE/DOCKET NUMBER: TAKAKURA-1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 628-5197
/ TELEFAX: (202) 737-3528
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 659 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-894-818B-1

Query Match
Best Local Similarity 20.4%; Score 457.5; DB 3; Length 659;
Matches 136; Conservative 69; Mismatches 153; Indels 101; Gaps 17;

QY 8 VKADVAGSSYGLYGGQIYVAVDTGLDGRNDSNHEAFRGKITALY-ALGRTNNANDTN 66
DB 145 IGADVTNLSIGYDGGVVAIVDTGIDAN-----HPDLKGVLTGMVDVNGSTPYDDQ 198
QY 67 GGTHTVAGSTLGNKATNK---GMAPOANLVFOSIM--DSSGGLGGLPSNLOTFSQAFSA 121
DB 199 GGTHTVAGSTLGNKATNK---GMAPOANLVFOSIM--DSSGGLGGLPSNLOTFSQAFSA 258
QY 122 GARI-----HTNSWGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPNGGT 169
DB 259 GIRVYNLSLSSQSSDGTBLSCAVNNAMD-----GIYVCVAAAGSGNTYT 306
QY 170 ISAPGTAKXAIITYGATENLRPSFGSYADNINNHVAFSSRGPTDGRIKPDVMAFGTYIIS 229
DB 307 VGSPAAASKVITVGA-----VDSNDNIAFSSSGPTADGRLKEVVAAGVDIIA 355
QY 230 AASSIAPDSSFFMNHDSKYAMGTSMATPIVAG--NVAQLREHFVNKRGITPK--PSLIK 286
DB 356 PRAS---GTSMGPINDIYTKAGTSMATPHVSGVALLIQAH-----PSWTDXK 404
QY 287 AALIAGA-----ADVGLGYPNGQMGKRVTLDKSL--NVAYNNESSALSTSKATY 335
DB 405 TALIETADIVAPKEIIDIAVGA-----GRVAVYAIKXVDYAKLTFFGSVADKGSATH 457
QY 336 TPTATAGPKLISLVMSDAPASTTASVTIVNDLITAPNGTRYGNDPFSAPFDNNMG 395
DB 458 TPTVSGATFTATILYND-----TGSSDIDLIVPNQME-----VDISYTA 498
QY 396 RNNVENYFINSFGSYTTEVQAVNPVGPNSLAIVN 434
DB 499 YGGEKVGYYNPTAGTWTVAVSYK---GAANYQVDVVS 534

RESULT 15
US-09-445-472-12
/ Sequence 12, Application US/09445472
/ Patent No. 6358726
/ GENERAL INFORMATION:
/ APPLICANT: TAKAKURA, Hikaru
/ APPLICANT: MORISHITA, Mio
/ APPLICANT: SHIMOTO, Tomoko
/ APPLICANT: ASADA, Kiyozo
/ APPLICANT: KATO, Ikunoshin
/ TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
/ FILE REFERENCE: TAKAKURA-6
/ CURRENT APPLICATION NUMBER: US/09/445,472
/ FILING DATE: 1999-12-06
/ PRIOR APPLICATION NUMBER: 151969/1997
/ PRIOR FILING DATE: 1997-06-10
/ NUMBER OF SEQ ID NOS: 33
/ SOFTWARE: Patent In version 3.0
/ SEQ ID NO 12
/ LENGTH: 659
/ TYPE: PRT
/ ORGANISM: Thermococcus celer
US-09-445-472-12
```

Sat Apr 3 17:19:59 2004

us-09-985-689a-2.rai

Page 10

Query Match	20.4%;	Score 457.5;	DB 4;	Length 659;
Best Local Similarity	29.6%;	Pred. No. 3,4e-29;		
Matches 136;	Conservative 69;	Mismatches 155;	Indels 101;	Gaps 17
QY	8	VKADVAOSSTGLVGQCIIVAAVDLTGLDGRNDSMHEAFRKITALY-ALGRTNANDTN	66	
Db	145	IGADIVYNSISGYDSSGVVAIVDTGIDAN-----HPDLKTKIGWYDANVGRSTPEYDQ	198	
QY	67	GHAHTVAGSVLVGNGAGNKK--GMAPQANVFGOSIM--DSSGGGLGGLPSNLQTFPGQAFSA	121	
Db	199	GHGHTVAGIVAGSGVNSQYIGVAGAKLVGKYVLGADGSSVSFTIAGVDWVQNKDKY	258	
QY	122	GARI-----HTNSWGAANVAGAYTDSRNVDYVRKNDMTLLFAAGNEGPNGT	169	
Db	259	GIRINILSLSSQSSSDGTSLQAVNNAMD-----GIIVCVAAAGSGPNFTY	306	
QY	170	ISAPGTAKNAITVAGDENLRPFSGVADINVAQFSSRGPTGGRIRKPYMAPGYIIS	229	
Db	307	VGSPAAASKIITVGA-----VDSNDIASFSSRGPTADGRIRKPEVMAPGVDTIA	355	
QY	230	ARSSLADSSFWANHDSKAYMGGTSMATPIVAG-NVAQLREHFVYKRGITPK--PSLTK	286	
Db	356	PRAS--GTSMGPIPDYVTKASGISMATPHVSGVGLITLQAF-----PSMTDPKVK	404	
QY	287	AALTAGA-----ADVGLGYPNNGQMGGRVTLKSL--NVAVYVSSALTSQKARY	335	
Db	405	TALLETADIVAPKEIDIAVGA-----GRVNVKAIKIDYAKLFFTSVSDVKDSATH	457	
QY	336	TFTATACKPIKISLVMSDAPASTTASVTLVNDLIVITAPNGTRYVGNDSAPFDNNMDG	395	
Db	458	TFDVSGATFTVTLALYMD-----TGSSDIDLYLDPNGNE-----VDYSYA	498	
QY	396	RNVENVFINSPOSQTYTIEVOAYNVPVQSPQSFSLATYN	434	
Db	499	YVGEKKGYINPFTAGTIVKVSYK--GAANYQVDVVS	534	

Search completed: March 31, 2004, 16:11:52
Job time : 15.0215 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:09:09 ; Search time 33.2177 Seconds

(without alignments)
3418.697 Million cell updates/sec

Title: US-09-985-689a-2

Perfect score: 2245
Sequence: 1 NDVARGIVKADVAQSSGYL.....EVQAINVEGPNFSLATVN 434

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2245	100.0	434	10	US-09-985-689a-2
2	2191	97.6	434	10	US-09-985-689a-1
3	2191	97.6	434	15	US-10-385-662-2
4	2130	94.9	434	10	US-09-985-689a-6
5	2118.5	94.4	434	10	US-09-985-689a-7
6	2010.5	89.6	433	10	US-09-985-689a-5
7	2006.5	89.4	433	10	US-09-985-689a-3
8	1999.5	89.1	433	10	US-09-985-689a-4
9	1579.5	70.4	345	14	US-10-336-324-10
10	1579.5	70.4	345	14	US-10-403-105-13
11	457.5	20.4	659	13	US-10-090-624-12
12	423	18.8	412	13	US-10-090-624-1
13	423	18.8	522	13	US-10-090-624-4
14	423	18.8	654	13	US-10-090-624-16
15	365	16.3	1079	14	US-10-112-488-39

16	360.5	16.1	1208	14	US-10-156-761-13251	Sequence 13251, A
17	345	15.4	1139	14	US-10-156-761-10856	Sequence 10856, A
18	343	15.3	1237	14	US-10-314-657-4	Sequence 4, App1
19	308.5	13.7	519	15	US-10-084-846A-114	Sequence 114, App
20	308.5	13.7	1975	15	US-10-084-846A-4	Sequence 59, App1
21	304.5	13.6	1388	13	US-10-090-624-6	Sequence 6, App1
22	281.5	12.5	595	10	US-09-927-827-59	Sequence 55, App1
23	277	12.3	580	10	US-09-927-827-55	Sequence 55, App1
24	273	12.2	418	9	US-09-966-921A-2	Sequence 12934, A
25	270	12.0	1101	14	US-10-156-761-12934	Sequence 255271, A
26	266	11.8	663	12	US-10-424-549-255271	Sequence 2, App1
27	256	11.4	280	14	US-10-209-812-9-2	Sequence 5, App1
28	256	11.4	397	10	US-09-779-334A-5	Sequence 3, App1
29	254	11.3	368	12	US-10-344-231-3	Sequence 2, App1
30	253	11.3	271	10	US-09-813-408-2	Sequence 56, App1
31	252	11.2	271	14	US-10-242-549-56	Sequence 6, App1
32	251.5	11.2	379	10	US-09-813-408-6	Sequence 52, App1
33	250	11.1	271	14	US-10-242-549-52	Sequence 46, App1
34	250	11.1	271	14	US-10-242-549-54	Sequence 48, App1
35	247	11.0	271	14	US-10-242-549-46	Sequence 60, App1
36	247	11.0	271	14	US-10-242-549-48	Sequence 16, App1
37	247	11.0	271	14	US-10-242-549-50	Sequence 1, App1
38	247	11.0	271	14	US-10-242-549-60	Sequence 8, App1
39	246	11.0	269	8	US-08-322-678-10	Sequence 49, App1
40	246	11.0	269	9	US-09-837-235-16	Sequence 1, App1
41	246	11.0	269	9	US-09-854-854B-6	Sequence 16, App1
42	246	11.0	269	9	US-09-575-139-1	Sequence 6, App1
43	246	11.0	269	9	US-09-576-414-8	Sequence 49, App1
44	246	11.0	269	9	US-09-736-116-49	Sequence 1, App1
45	246	11.0	269	13	US-10-075-907-1	Sequence 1, App1

ALIGNMENTS

RESULT 1
US-09-985-689a-2
Sequence 2, Application US/09985689a
Publication No. US2003002235:1A1
GENERAL INFORMATION:
APPLICANT: HATADA, YUJI
APPLICANT: OGAWA, AKINORI
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SATO, TSUYOSHI
APPLICANT: ARAKI, HIROYUKI
APPLICANT: SUMITOMO, NOBUYUKI
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAKET, KATSUHIKA
TITLE OR INVENTION: Alkaline proteases
FILE REFERENCE: 215483US0
CURRENT APPLICATION NUMBER: US/09/985.689a
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: JP P2000-355166
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: JP P2001-114048
PRIOR FILING DATE: 2001-04-12
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 434
TYPE: PRT
ORGANISM: Bacillus sp.
US-09-985-689a-2
Query Match 100.0%; Score 2245; DB 10; Length 434;
Best Local Similarity 100.0%; Pred. No. 8.4e-189;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQSSGYLGGQGIYAVADTGIDTGRNDSWHEAFRGKITLALAGRTN 60
DB 1 NDVARGIVKADVAQSSGYLGGQGIYAVADTGIDTGRNDSWHEAFRGKITLALAGRTN 60
QY 61 NADTNGHGTIVAGSVLNGATNKKMAPOANTVFOSIMDSGGLGLPENTLQTLFSGQAFS 120

```

Db      61 NANDTNGHGVASVGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLQTLFSGQAFS 120
Qy      121 AGARIHNSWGAANVNGAYTTDSRVDDYVAKNDMTILFAAGNEGPNGGTISAPGTAKNAI 180
Db      121 AGARIHNSWGAANVNGAYTTDSRVDDYVAKNDMTILFAAGNEGPNGGTISAPGTAKNAI 180
Qy      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 240
Db      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 240
Qy      241 WANHSKYAVYGGSMATPIVAGNVAQLREHFVNKRGITPKPSLLKALLAGAADVGLGY 300
Db      241 WANHSKYAVYGGSMATPIVAGNVAQLREHFVNKRGITPKPSLLKALLAGAADVGLGY 300
Qy      301 PNGNQGWRVTLDKSLNVAAYNESSALSTGKATYFTATAGKPKLSIWSMDAPASTTA 360
Db      301 PNGNQGWRVTLDKSLNVAAYNESSALSTGKATYFTATAGKPKLSIWSMDAPASTTA 360
Qy      361 SVTLVNDLDLVITAPNGTRYVGNDFSAPEDNMDGRNNEVNFINSQSGTYITIEVOAYN 420
Db      361 SVTLVNDLDLVITAPNGTRYVGNDFSAPEDNMDGRNNEVNFINSQSGTYITIEVOAYN 420
Qy      421 VPVGPQNFSLAIYN 434
Db      421 VPVGPQNFSLAIYN 434

```

RESULT 2

```

US-09-985-689a-1
; Sequence 1, Application US/09985689a
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483U90
; CURRENT APPLICATION NUMBER: US/09/985,689a
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689a-1

```

```

Query Match      97.6%; Score 2191; DB 10; Length 434;
Best Local Similarity 96.5%; Pred. No. 4.7e-184;
Matches 419; Conservative 13; Mismatches 2; Indels 0; Gaps 0;
Qy      1 NDVARGIVKADVAQSSVGLYGGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRN 60
Db      1 NDVARGIVKADVAQSSVGLYGGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRN 60
Qy      61 NANDTNGHGVASVGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLQTLFSGQAFS 120
Db      61 NANDTNGHGVASVGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLQTLFSGQAFS 120
Qy      121 AGARIHNSWGAANVNGAYTTDSRVDDYVAKNDMTILFAAGNEGPNGGTISAPGTAKNAI 180
Db      121 AGARIHNSWGAANVNGAYTTDSRVDDYVAKNDMTILFAAGNEGPNGGTISAPGTAKNAI 180
Qy      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 240
Db      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 240

```

```

Qy      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 240
Db      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 240
Qy      241 WANHSKYAVYGGSMATPIVAGNVAQLREHFVNKRGITPKPSLLKALLAGAADVGLGY 300
Db      241 WANHSKYAVYGGSMATPIVAGNVAQLREHFVNKRGITPKPSLLKALLAGAADVGLGY 300
Qy      301 PNGNQGWRVTLDKSLNVAAYNESSALSTGKATYFTATAGKPKLSIWSMDAPASTTA 360
Db      301 PNGNQGWRVTLDKSLNVAAYNESSALSTGKATYFTATAGKPKLSIWSMDAPASTTA 360
Qy      361 SVTLVNDLDLVITAPNGTRYVGNDFSAPEDNMDGRNNEVNFINSQSGTYITIEVOAYN 420
Db      361 SVTLVNDLDLVITAPNGTRYVGNDFSAPEDNMDGRNNEVNFINSQSGTYITIEVOAYN 420
Qy      421 VPVGPQNFSLAIYN 434
Db      421 VPVGPQNFSLAIYN 434

```

RESULT 3

```

US-10-385-662-2
; Sequence 2, Application US/10385662
; Publication No. US20040002432A1
; GENERAL INFORMATION:
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: SAITO, KAZUHIRO
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: IZAWA, YOSHIFUMI
; APPLICANT: SAKKI, KATSUHIISA
; APPLICANT: KOBAYASHI, TOHRU
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: Alkaline protease
; FILE REFERENCE: 234936U90
; CURRENT APPLICATION NUMBER: US/10/385,662
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: JP 2002-081428
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: JP 2002-165987
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: JP 2002-304230
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: JP 2002-304231
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp. KSM-KP43
US-10-385-662-2

```

```

Query Match      97.6%; Score 2191; DB 15; Length 434;
Best Local Similarity 96.5%; Pred. No. 4.7e-184;
Matches 419; Conservative 13; Mismatches 2; Indels 0; Gaps 0;
Qy      1 NDVARGIVKADVAQSSVGLYGGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRN 60
Db      1 NDVARGIVKADVAQSSVGLYGGQGIIVAVADTGLDTGRNDSMHEAFRGKITLALYALGRN 60
Qy      61 NANDTNGHGVASVGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLQTLFSGQAFS 120
Db      61 NANDTNGHGVASVGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLQTLFSGQAFS 120
Qy      121 AGARIHNSWGAANVNGAYTTDSRVDDYVAKNDMTILFAAGNEGPNGGTISAPGTAKNAI 180
Db      121 AGARIHNSWGAANVNGAYTTDSRVDDYVAKNDMTILFAAGNEGPNGGTISAPGTAKNAI 180
Qy      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 240
Db      181 TVGATEMLRPSFGSYADININHVAFSSRGPTKGRIKPDVMAFGTYILSARSSLAPDSSF 240

```

```

QY 241 WANDSKYAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVGLGY 300
DB 241 WANDSKYAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVGLGY 300
QY 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYTTAGKPLKISLWSDAPASTTA 360
DB 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYTTAGKPLKISLWSDAPASTTA 360
QY 361 SVTLVNDLDLVTAPNGTRVYGNDFSAFPDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
DB 361 SVTLVNDLDLVTAPNGTRVYGNDFSAFPDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
QY 421 VPVGPONFSLATVN 434
DB 421 VPVGPONFSLATVN 434

```

RESULT 4

```

US-09-985-689a-6
; Sequence 6, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 434
; TYPE: PRP
; ORGANISM: Bacillus sp.
US-09-985-689a-6

```

Query Match 94.4%; Score 2130; DB 10; Length 434;

Best Local Similarity 92.4%; Pred. No. 1, 1e-178; Index 0; Gaps 0;

```

QY 1 NDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMEAFRGKITLALYALGRTN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMEAFRGKITLALYALGRTN 60
QY 61 NANDTNGHGHVAGSVYGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLQTLFSGAYS 120
DB 61 NANDTNGHGHVAGSVYGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLQTLFSGAYS 120
QY 121 AGARHITNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGPNNGTISAPGTAKNAI 180
DB 121 AGARHITNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGPNNGTISAPGTAKNAI 180
QY 181 TVGATEMLRPSFGSYADNINHVAFSSRGPTRKGRIPDYMAPGTITLSARSLAPDSST 240
DB 181 TVGATEMLRPSFGSYADNINHVAFSSRGPTRKGRIPDYMAPGTITLSARSLAPDSST 240
QY 241 WANDSKYAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVGLGY 300
DB 241 WANDSKYAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVGLGY 300
QY 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYTTAGKPLKISLWSDAPASTTA 360
DB 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYTTAGKPLKISLWSDAPASTTA 360

```

```

DB 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYTTAGKPLKISLWSDAPASTTA 360
QY 361 SVTLVNDLDLVTAPNGTRVYGNDFSAFPDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
DB 361 SVTLVNDLDLVTAPNGTRVYGNDFSAFPDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
QY 421 VPVGPONFSLATVN 434
DB 421 VPVGPONFSLATVN 434

```

RESULT 5

```

US-09-985-689a-7
; Sequence 7, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 433
; TYPE: PRP
; ORGANISM: Bacillus sp.
US-09-985-689a-7

```

Query Match 94.4%; Score 2118.5; DB 10; Length 433;

Best Local Similarity 93.3%; Pred. No. 1, 1e-177; Index 1; Gaps 1;

```

QY 1 NDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMEAFRGKITLALYALGRTN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGIYAVADTGLDTRGNDSSMEAFRGKITLALYALGRTN 60
QY 61 NANDTNGHGHVAGSVYGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLQTLFSGAYS 120
DB 61 NANDTNGHGHVAGSVYGNAGATNKGMAPQANLVFQSIIMDSGGIGLPSNLQTLFSGAYS 120
QY 121 AGARHITNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGPNNGTISAPGTAKNAI 180
DB 121 AGARHITNSWGAAYNGAYTTDSRNVDYVRKNDMTILFAAGNEGPNNGTISAPGTAKNAI 180
QY 181 TVGATEMLRPSFGSYADNINHVAFSSRGPTRKGRIPDYMAPGTITLSARSLAPDSST 240
DB 181 TVGATEMLRPSFGSYADNINHVAFSSRGPTRKGRIPDYMAPGTITLSARSLAPDSST 240
QY 241 WANDSKYAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVGLGY 300
DB 241 WANDSKYAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSLLKALILAGADVGLGY 300
QY 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYTTAGKPLKISLWSDAPASTTA 360
DB 301 PNGNGGGRVTLDKSLNVAAYNNESSALSTOKATYTTAGKPLKISLWSDAPASTTA 360
QY 361 SVTLVNDLDLVTAPNGTRVYGNDFSAFPDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
DB 361 SVTLVNDLDLVTAPNGTRVYGNDFSAFPDNNMDGRNNVENVFINSPOSGTYTIEVOAYN 420
QY 421 VPVGPONFSLATVN 434
DB 421 VPVGPONFSLATVN 434

```

Db 420 VPVSPQRFSLAIHV 433

RESULT 6

US-09-985-689A-5
 ; Sequence 5, Application US/09985689A
 ; Publication No. US2003022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAWA, AKINORI
 ; APPLICANT: KAGEYAMA, YASUSHI
 ; APPLICANT: SAKI, KATSUHIISA
 ; APPLICANT: SATO, TSIYOSHI
 ; APPLICANT: ARAKI, HIROYUKI
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: OKUDA, MITSUYOSHI
 ; APPLICANT: SAKI, KATSUHIISA
 ; TITLE OF INVENTION: Alkaline proteases
 ; FILE REFERENCE: 215483US0
 ; CURRENT APPLICATION NUMBER: US/09/985,689A
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: JP P2000-355166
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: JP P2001-114048
 ; PRIOR FILING DATE: 2001-04-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 433
 ; TYPE: PR1
 ; ORGANISM: Bacillus sp.
 US-09-985-689A-5

Query Match 89.6%; Score 2010.5; DB 10; Length 433;
 Best Local Similarity 88.5%; Pred. No. 3,6e-168;
 Matches 384; Conservative 26; Mismatches 23; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMEAFRGKITALYALGRTN 60
 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMEAFRGKITALYALGRTN 60
 Db 61 NANTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMSSGGIGLPSNLQTLFSGAQS 120
 61 NANTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMSSGGIGLPSNLQTLFSGAQS 120
 Db 61 NANTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMSSGGIGLPSNLQTLFSGAQS 119
 61 NANTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMSSGGIGLPSNLQTLFSGAQS 119
 QY 121 AGARIHTNSWGAANGAAYTTDSRVDDYRKNDMTILFAAGNEGPNGGTTISAPGTAKNAI 180
 121 AGARIHTNSWGAANGAAYTTDSRVDDYRKNDMTILFAAGNEGPNGGTTISAPGTAKNAI 180
 Db 120 AGARIHTNSWGAANGAAYTTDSRVDDYRKNDMTILFAAGNEGPNGGTTISAPGTAKNAI 179
 120 AGARIHTNSWGAANGAAYTTDSRVDDYRKNDMTILFAAGNEGPNGGTTISAPGTAKNAI 179
 QY 181 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQTYTILSARSSLADDSF 240
 181 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQTYTILSARSSLADDSF 240
 Db 180 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQTYTILSARSSLADDSF 239
 180 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQTYTILSARSSLADDSF 239
 QY 241 WANHSKXAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSILKALILAGADYGLGY 300
 241 WANHSKXAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSILKALILAGADYGLGY 300
 Db 240 WANHSKXAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSILKALILAGADYGLGY 299
 240 WANHSKXAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSILKALILAGADYGLGY 299
 QY 301 PNGQGWGRVTLDSKSLNVAAYNESSALSTSOKATYTTATAGAPLKISLWSDAPASTTA 360
 301 PNGQGWGRVTLDSKSLNVAAYNESSALSTSOKATYTTATAGAPLKISLWSDAPASTTA 360
 Db 300 PNGQGWGRVTLDSKSLNVAAYNESSALSTSOKATYTTATAGAPLKISLWSDAPASTTA 359
 300 PNGQGWGRVTLDSKSLNVAAYNESSALSTSOKATYTTATAGAPLKISLWSDAPASTTA 359
 QY 361 SVTLVNDLIDLVITAPNGRTYVGNDFSAFPDNNMDGRNNVENVFINSQSGTYTIEVOAYN 420
 361 SVTLVNDLIDLVITAPNGRTYVGNDFSAFPDNNMDGRNNVENVFINSQSGTYTIEVOAYN 420
 Db 360 SVTLVNDLIDLVITAPNGRTYVGNDFSAFPDNNMDGRNNVENVFINSQSGTYTIEVOAYN 419
 360 SVTLVNDLIDLVITAPNGRTYVGNDFSAFPDNNMDGRNNVENVFINSQSGTYTIEVOAYN 419
 QY 421 VPVGPQNFSLAIYN 434
 421 VPVGPQNFSLAIYN 434
 Db 420 VPVGPQNFSLAIHV 433

RESULT 7
 US-09-985-689A-3
 ; Sequence 3, Application US/09985689A

Publication No. US2003022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAWA, AKINORI
 ; APPLICANT: KAGEYAMA, YASUSHI
 ; APPLICANT: SAKI, KATSUHIISA
 ; APPLICANT: SATO, TSIYOSHI
 ; APPLICANT: ARAKI, HIROYUKI
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: OKUDA, MITSUYOSHI
 ; APPLICANT: SAKI, KATSUHIISA
 ; TITLE OF INVENTION: Alkaline proteases
 ; FILE REFERENCE: 215483US0
 ; CURRENT APPLICATION NUMBER: US/09/985,689A
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: JP P2000-355166
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: JP P2001-114048
 ; PRIOR FILING DATE: 2001-04-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 433
 ; TYPE: PR1
 ; ORGANISM: Bacillus sp.
 US-09-985-689A-3

Query Match 89.4%; Score 2006.5; DB 10; Length 433;
 Best Local Similarity 88.2%; Pred. No. 8e-168;
 Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMEAFRGKITALYALGRTN 60
 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRGNDSSMEAFRGKITALYALGRTN 60
 Db 61 NANTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMSSGGIGLPSNLQTLFSGAQS 120
 61 NANTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMSSGGIGLPSNLQTLFSGAQS 120
 Db 61 NANTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMSSGGIGLPSNLQTLFSGAQS 119
 61 NANTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMSSGGIGLPSNLQTLFSGAQS 119
 QY 121 AGARIHTNSWGAANGAAYTTDSRVDDYRKNDMTILFAAGNEGPNGGTTISAPGTAKNAI 180
 121 AGARIHTNSWGAANGAAYTTDSRVDDYRKNDMTILFAAGNEGPNGGTTISAPGTAKNAI 180
 Db 120 AGARIHTNSWGAANGAAYTTDSRVDDYRKNDMTILFAAGNEGPNGGTTISAPGTAKNAI 179
 120 AGARIHTNSWGAANGAAYTTDSRVDDYRKNDMTILFAAGNEGPNGGTTISAPGTAKNAI 179
 QY 181 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQTYTILSARSSLADDSF 240
 181 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQTYTILSARSSLADDSF 240
 Db 180 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQTYTILSARSSLADDSF 239
 180 TVGATEENLRPSFGSVADNINHVAFSSRGPTKGRIRKPDVMAQTYTILSARSSLADDSF 239
 QY 241 WANHSKXAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSILKALILAGADYGLGY 300
 241 WANHSKXAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSILKALILAGADYGLGY 300
 Db 240 WANHSKXAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSILKALILAGADYGLGY 299
 240 WANHSKXAYMGTSMTPTVAGNVAQLREHFVKNRGITPKPSILKALILAGADYGLGY 299
 QY 301 PNGQGWGRVTLDSKSLNVAAYNESSALSTSOKATYTTATAGAPLKISLWSDAPASTTA 360
 301 PNGQGWGRVTLDSKSLNVAAYNESSALSTSOKATYTTATAGAPLKISLWSDAPASTTA 360
 Db 300 PNGQGWGRVTLDSKSLNVAAYNESSALSTSOKATYTTATAGAPLKISLWSDAPASTTA 359
 300 PNGQGWGRVTLDSKSLNVAAYNESSALSTSOKATYTTATAGAPLKISLWSDAPASTTA 359
 QY 361 SVTLVNDLIDLVITAPNGRTYVGNDFSAFPDNNMDGRNNVENVFINSQSGTYTIEVOAYN 420
 361 SVTLVNDLIDLVITAPNGRTYVGNDFSAFPDNNMDGRNNVENVFINSQSGTYTIEVOAYN 420
 Db 360 SVTLVNDLIDLVITAPNGRTYVGNDFSAFPDNNMDGRNNVENVFINSQSGTYTIEVOAYN 419
 360 SVTLVNDLIDLVITAPNGRTYVGNDFSAFPDNNMDGRNNVENVFINSQSGTYTIEVOAYN 419
 QY 421 VPVGPQNFSLAIYN 434
 421 VPVGPQNFSLAIYN 434
 Db 420 VPVGPQNFSLAIHV 433

RESULT 8

US-09-985-689A-4
 ; Sequence 4, Application US/09985689A
 ; Publication No. US2003022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAWA, AKINORI
 ; APPLICANT: KAGEYAMA, YASUSHI
 ; APPLICANT: SAKI, KATSUHIISA
 ; APPLICANT: SATO, TSIYOSHI
 ; APPLICANT: ARAKI, HIROYUKI

APPLICANT: SUMITOMO, NOBUYUKI
 APPLICANT: OKUDA, MITSUYOSHI
 APPLICANT: SAKI, KATSUHIISA
 TITLE OF INVENTION: Alkaline proteases
 FILE REFERENCE: 215483USO
 CURRENT APPLICATION NUMBER: US/09/985,689A
 PRIOR FILING DATE: 2002-07-01
 PRIOR APPLICATION NUMBER: JP P2000-355166
 PRIOR FILING DATE: 2000-11-22
 PRIOR APPLICATION NUMBER: JP P2001-114048
 PRIOR FILING DATE: 2001-04-12
 NUMBER OF SEQ ID NOS: 7
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 4
 LENGTH: 433
 TYPE: PRT
 ORGANISM: Bacillus sp.
 US-09-985-689A-4

Query Match 89.1%; Score 1999.5; DB 10; Length 433;
 Best Local Similarity 86.0%; Pred. No. 3.3e-167;
 Matches 362; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTRNDSMHEAFRGKITALVALGRTN 60
 1 NDVARGIVADVAQNNYGLYGQGVAVADTGLDTRNDSMHEAFRGKITALVALGRTN 60
 QY 61 NANDNGHGHVAGSVLGNATNKGAPOANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 120
 61 NANDNGHGHVAGSVLGN-ALNKGAPOANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 119
 DB 121 AGARIHTNSWGAANGAYTTDSRNVDYRKNDMTLLFAAGNEGPNGTISAPGTAKNAI 180
 120 AGARIHTNSWGAANGAYTTDSRNVDYRKNDMTLLFAAGNEGPNGTISAPGTAKNAI 179
 DB 181 TVGATENTLRPSFSGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTYILSARSSLADPDSF 240
 180 TVGATENTLRPSFSGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTYILSARSSLADPDSF 239
 DB 241 WANDSKYAVMGSTMAPIVAGNVAQLREHFVKRGTIPKPSLLKALILAGADVGLGF 300
 240 WANDSKYAVMGSTMAPIVAGNVAQLREHFVKRGTIPKPSLLKALILAGADVGLGF 299
 QY 301 PNGNQGWGRTVLDKSLNV 318
 300 PNGNQGWGRTVLDKSLNV 317
 DB 361 SVTLVVDLVLVTAPNGTRYVGNDFSAFPDNNMDGNNVENVFINSPOSGTYTIEVOAYN 420
 360 SVTLVVDLVLVTAPNGTRYVGNDFSAFPDNNMDGNNVENVFINSPOSGTYTIEVOAYN 419
 QY 421 VVPGPQNFSLAIYN 434
 420 VVPGPQNFSLAIYN 433
 DB 420 VVPGPQNFSLAIYN 433

RESULT 9
 US-10-336-324-10
 Sequence 10; Application US/10336324
 Publication No. US20030176304A1
 GENERAL INFORMATION:
 APPLICANT: Hansen, Peter
 APPLICANT: Bauditz, Peter
 APPLICANT: Mikkelsen, Frank
 TITLE OF INVENTION: Protease Variants and Compositions
 FILE REFERENCE: 5349,204-US
 CURRENT APPLICATION NUMBER: US/10/336,324
 PRIOR FILING DATE: 2003-01-03
 PRIOR APPLICATION NUMBER: US/09/5,2,251A
 PRIOR FILING DATE: 2000-02-24
 NUMBER OF SEQ ID NOS: 12
 SOFTWARE: PatentIn version 3.1

SEQ ID NO 10
 LENGTH: 345
 TYPE: PRT
 ORGANISM: Bacillus
 US-10-336-324-10

Query Match 70.4%; Score 1579.5; DB 14; Length 345;
 Best Local Similarity 95.6%; Pred. No. 2.1e-130;
 Matches 304; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

QY 1 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTRNDSMHEAFRGKITALVALGRTN 60
 29 NDVARGIVADVAQNNYGLYGQGVAVADTGLDTRNDSMHEAFRGKITALVALGRTN 88
 DB 61 NANDNGHGHVAGSVLGNATNKGAPOANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 120
 89 NANDNGHGHVAGSVLGN-ATNKGAPOANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 147
 QY 121 AGARIHTNSWGAANGAYTTDSRNVDYRKNDMTLLFAAGNEGPNGTISAPGTAKNAI 180
 148 AGARIHTNSWGAANGAYTTDSRNVDYRKNDMTLLFAAGNEGPNGTISAPGTAKNAI 207
 DB 208 TVGATENTLRPSFSGSYADNINHVAFSSRGPTKDGRIKPDVMAFGTYILSARSSLADPDSF 267
 241 WANDSKYAVMGSTMAPIVAGNVAQLREHFVKRGTIPKPSLLKALILAGADVGLGF 300
 268 WANDSKYAVMGSTMAPIVAGNVAQLREHFVKRGTIPKPSLLKALILAGADVGLGF 327
 QY 301 PNGNQGWGRTVLDKSLNV 318
 328 PNGNQGWGRTVLDKSLNV 345
 DB 328 PNGNQGWGRTVLDKSLNV 345

RESULT 10
 US-10-403-105-13
 Sequence 13; Application US/10403105
 Publication No. US20030180933A1
 GENERAL INFORMATION:
 APPLICANT: Hansen, Peter K.
 APPLICANT: Bauditz, Peter
 APPLICANT: Mikkelsen, Frank
 TITLE OF INVENTION: Protease Variants And Compositions
 FILE REFERENCE: 5435,200-US
 CURRENT APPLICATION NUMBER: US/10/403,105
 CURRENT FILING DATE: 2003-03-31
 PRIOR APPLICATION NUMBER: US/09/196,261A
 PRIOR FILING DATE: 1998-11-19
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1332/97
 PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-21
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 13
 LENGTH: 345
 TYPE: PRT
 ORGANISM: Bacillus
 US-10-403-105-13

Query Match 70.4%; Score 1579.5; DB 14; Length 345;
 Best Local Similarity 95.6%; Pred. No. 2.1e-130;
 Matches 304; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

QY 1 NDVARGIVADVAQSSYGLYGQGIYAVADTGLDTRNDSMHEAFRGKITALVALGRTN 60
 29 NDVARGIVADVAQNNYGLYGQGVAVADTGLDTRNDSMHEAFRGKITALVALGRTN 88
 DB 61 NANDNGHGHVAGSVLGNATNKGAPOANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 120
 89 NANDNGHGHVAGSVLGN-ATNKGAPOANLVFQSIIMSSGGLGGLPSNLQTLFSGQAFS 147
 QY 121 AGARIHTNSWGAANGAYTTDSRNVDYRKNDMTLLFAAGNEGPNGTISAPGTAKNAI 180

Sat Apr 3 17:19:59 2004

us-09-985-689a-2.rapb

Page 6

Db 148 AGARHITNSGAFVNGAYTTDSRNVDDYRKNDMTLLFAAGNEGPGSGTISAPGTAKNAI 207
Qy 181 TVGATENLRPSFGSYADININHVAFSSRGPTKGRIRKIPVMAFGYITLSARSLAPDSF 240
Db 208 TVGATENLRPSFGSYADININHVAFSSRGPTKGRIRKIPVMAFGYITLSARSLAPDSF 267
Qy 241 WANHSKYAYMGSTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKALIAADVGIGY 300
Db 268 WANHSKYAYMGSTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKALIAADVGIGY 327
Qy 301 PNGNGGRTYLDKSLNV 318
Db 328 PNGNGGRTYLDKSLNV 345

RESULT 11
US-10-090-624-12
; Sequence 12, Application US/10090624
; Publication No. US2002013235A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match 20.4%; Score 457.5; DB 13; Length 659;
Best Local Similarity 29.6%; Pred. No. 2,7e-31;
Matches 136; Conservative 69; Mismatches 153; Indels 101; Gaps 17;

Qy 8 VKADVAQSYGLYGGQIVAVADTGLDTRNDSSMEAFRGKITALY-ALGRTNANDTN 66
Db 145 IGAATVANSLSGYDSSGVVAIVDTGITAN-----HDLKGRYIGMTDAVNGSTFYDQ 198
Qy 67 GHGHTVAGSVLGNATNK---GMAPQANLVFQSIM--DSSGGLGGLPSNLQTLFSGAFBA 121
Db 199 GHGHTVAGSVLGNATNK---GMAPQANLVFQSIM--DSSGGLGGLPSNLQTLFSGAFBA 258
Qy 122 GARI-----HTNSGAAVNGAYTTDSRNVDDYRKNDMTLLFAAGNEGNGGT 169
Db 259 GIRIINSLGSSGSDGTDLSQAIVNANMDA-----GIYCVAAAGSGRIT 306
Qy 170 ISAGGTAKNAITVATENLRPSFGSYADININHVAFSSRGPTKGRIRKIPVMAFGYITLS 229
Db 307 VGSFAAASKVITVGA-----VDSNDNLSFSSRGPTADGRIRKIPVMAFGYITLS 355
Qy 230 ARSLADSSFWANHSKYAYMGSTSVATPIVAG-NAQLREHFVKNRGITPK--PSLLK 286
Db 356 PRAS---GTSWGTINDYITKASGTSMATPHVAGSVGLIQAQ-----PSWTPDKRK 404
Qy 287 AALIGA-----ADVGLGYPNGNGGRTYLDKSL---NVAAYNESSALSTQKATY 335
Db 405 TALJETADIYAPKEIADIYGA-----GRVNYKAIKYDDYAKLFTGSVADKGSATH 457
Qy 336 TETATACKPLKISLVMSDAPASTASVTLVNDLDTVITANGRIYVNDSSAFDDNWDG 395
Db 458 TFDVSGATFVATLYMD-----TGSDDLDLYLDPVNGE-----VDYSYTA 498

Qy 396 RNNVENFINSPOSITIEVCANVPEPONESLAIYN 434
Db 499 YGGEKVGYNPFTAGTWTYKVASYK---GAANYQVDVVS 534

RESULT 12
US-10-090-624-1
; Sequence 1, Application US/10090624
; Publication No. US2002013235A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-10-090-624-1

Query Match 18.8%; Score 423; DB 13; Length 412;
Best Local Similarity 29.6%; Pred. No. 1,5e-28;
Matches 135; Conservative 61; Mismatches 146; Indels 114; Gaps 18;

Qy 18 GLYGGQIVAVADTGLDTRNDSSMEAFRGKITALYALGRTNAN-----DTNGHGT 71
Db 22 GYDSGIGITIGITIDID-----ASHPDLOGKV-----IGWPFVNGRSYDYDHGHT 70
Qy 72 VAGSVLGNATNK---GMAPQANLVFQSIM--DSSGGLGGLPSNLQTLFSGAFBA 125
Db 71 VASIAAGTGAASNGKRYGMAPQALAGIKVLAGDSSISITIKGYMAVANDKDKIKY 130
Qy 126 HTNSGA-----AVNGAYTTDSRNVDDYRKNDMTLLFAAGNEGNGGTISAP 173
Db 131 INLSGSSGSDGTDLSQAIVNANMDA-----GLVYVAAAGSGPKNYITIGSP 178
Qy 174 GTAKNAITVATENLRPSFGSYADININHVAFSSRGPTKGRIRKIPVMAFGYITLSARSS 233
Db 179 AAASKVITVGA-----VSKYVITFSRGPTADGRIRKIPVMAFGYITLSARSS 227
Qy 234 LAPSSFWANHSKYAYMGSTSVATPIVAGNVAQLREHFVKNRGITPK--PSLLKALIA 291
Db 228 ---GTSWGTINDYITKASGTSMATPHVAGSVGLIQAQ-----AHSWTPDKRKALIE 277
Qy 292 GA-----ADVGLGYPNGNGGRTYLDKSLVAVNESSALSTQKATY 338
Db 278 TADIYKPEIADIYGA-----GRVNYKAIKYDDYAKLFTGSVADKGSATH 457
Qy 339 ATAGKPLKISLVMSDAPASTASVTLVNDLDTVITANGRIYVNDSSAFDDNWDG 395
Db 339 ISGASFVATLYMDAN-----SDLDLYYDENGNG-----VDYSYTA 498
Qy 399 VENNFINSPGSGTITIEVCANVPEPONESLAIYN 434
Db 370 FEKVGYNPFTAGTWTYKVASYK---GAANYQVDVVS 534

RESULT 13
US-10-090-624-4
; Sequence 4, Application US/10090624
; Publication No. US2002013235A1
; GENERAL INFORMATION:


```

; APPLICANT: TAKAKURA, HiKaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOUO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (428)..(428)
; OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-10-090-624-4

Query Match      18.8%; Score 423; DB 13; Length 522;
Best Local Similarity 29.6%; Pred. No. 2.1e-28;
Matches 135; Conservative 61; Mismatches 146; Indels 114; Gaps 18;

QY 18 GLYGQGIIVAVADTGLDTGRNDSMEAFRGKITALYALGRINNAN-----DTNGHGH 71
DB 22 GYDSSGITTIGITDID-----ASHPDLOGKV-----IGWDFVNGRSYPYDDHGHGH 70
QY 72 VAGSVLGNAGTN-----KGNAPQANLVFQSIM--DSSGGLGGLPSNLQTLFSGAFSAGARI 125
DB 71 VASIAAGTGAASNGKYGMAFGAKLAGIKVLGADSGSISTIKGVEMAVDNKKDKYGIKV 130
QY 126 HTSMWGA-----AVNGAYTTDSRNVDYVRKNDMTILFPAEGNGGTTISAP 173
DB 131 INSLGSSQSSDGTDLAQVNAAMD-----GLVVVAAGNSGPNKYYITGSP 178
QY 174 GTAKNATTVGATENTLRPSFGSYADNINHAQFSSRGTGRKPYMAFGTYILSARS 233
DB 179 AASKVITVGA-----VDKYDVITSFSSRGPTADGRKPEVVAFGNMTIAARS 227
QY 234 LAPDSSFWANHDSKYAVMGTSMTATPIVAGNVAQLREHFVKNRGITPK--PSLLKALIA 291
DB 228 --GTSMGQPIINDYTTAAPGTSMTATPIVAGIALLLQ-----AHPSWTPEKVKTALIE 277
QY 292 GA-----ADVGLGYPNNGQMGHVTLDKSLNVAYNSSALSTSOKA-----TYTFT 338
DB 278 TADIVKDEIADLIVYGA-----GRVNAVKAIN--YDNVAKLVFTGYVANKGSQTHQFV 328
QY 339 ATAGKPKLISLVWSDAPASTTASVTLVNDLVLITAPNGTRYVGNDFSAFPDNNMDGRNN 398
DB 329 ISGASFTATLYDMNAN-----SDLDLYLYDPNNGQ-----VDYSTAYYG 369
QY 399 VENVFINSPOSQTYTIEVQAVNVFVGPONFSLAIYN 434
DB 370 FEKVGYNPTDGTWTIKVYSYSS--GSANYQVYVVS 402

```

```

; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-10-090-624-16

Query Match      18.8%; Score 423; DB 13; Length 654;
Best Local Similarity 29.6%; Pred. No. 2.9e-28;
Matches 135; Conservative 61; Mismatches 146; Indels 114; Gaps 18;

QY 18 GLYGQGIIVAVADTGLDTGRNDSMEAFRGKITALYALGRINNAN-----DTNGHGH 71
DB 154 GYDSSGITTIGITDID-----ASHPDLOGKV-----IGWDFVNGRSYPYDDHGHGH 202
QY 72 VAGSVLGNAGTN-----KGNAPQANLVFQSIM--DSSGGLGGLPSNLQTLFSGAFSAGARI 125
DB 203 VASIAAGTGAASNGKYGMAFGAKLAGIKVLGADSGSISTIKGVEMAVDNKKDKYGIKV 262
QY 126 HTSMWGA-----AVNGAYTTDSRNVDYVRKNDMTILFPAEGNGGTTISAP 173
DB 263 INSLGSSQSSDGTDLAQVNAAMD-----GLVVVAAGNSGPNKYYITGSP 310
QY 174 GTAKNATTVGATENTLRPSFGSYADNINHAQFSSRGTGRKPYMAFGTYILSARS 233
DB 311 AASKVITVGA-----VDKYDVITSFSSRGPTADGRKPEVVAFGNMTIAARS 359
QY 234 LAPDSSFWANHDSKYAVMGTSMTATPIVAGNVAQLREHFVKNRGITPK--PSLLKALIA 291
DB 360 --GTSMGQPIINDYTTAAPGTSMTATPIVAGIALLLQ-----AHPSWTPEKVKTALIE 409
QY 292 GA-----ADVGLGYPNNGQMGHVTLDKSLNVAYNSSALSTSOKA-----TYTFT 338
DB 410 TADIVKDEIADLIVYGA-----GRVNAVKAIN--YDNVAKLVFTGYVANKGSQTHQFV 460
QY 339 ATAGKPKLISLVWSDAPASTTASVTLVNDLVLITAPNGTRYVGNDFSAFPDNNMDGRNN 398
DB 461 ISGASFTATLYDMNAN-----SDLDLYLYDPNNGQ-----VDYSTAYYG 501
QY 399 VENVFINSPOSQTYTIEVQAVNVFVGPONFSLAIYN 434
DB 502 FEKVGYNPTDGTWTIKVYSYSS--GSANYQVYVVS 534

```

```

RESULT 14
US-10-090-624-16
; Sequence 16, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, HiKaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOUO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE

```

```

RESULT 15
US-10-112-488-39
; Sequence 39, Application US/10112488
; Publication No. US20030082746A1
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, Yoshimi
; APPLICANT: UMEZAWA, Yukiko
; APPLICANT: YOKOYAMA, Keiichi
; APPLICANT: MATSUI, Hiroshi
; TITLE OF INVENTION: PROCESS FOR PRODUCING TRANSLUTAMINASE
; FILE REFERENCE: 219286USOCONT
; CURRENT APPLICATION NUMBER: US/10/112,488
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: PCT/JP00/06780
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: J22000-280098
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: JP11-280098
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 70

```

SOFTWARE: PatentIn version 3.1
SEQ ID NO 39
LENGTH: 1079
TYPE: PRT
ORGANISM: Streptomyces albobogisecius
US-10-112-488-39

Query Match 16.3%; Score 365; DB 14; Length 1079;
Best Local Similarity 31.3%; Pred. No. 7.7e-23;
Matches 146; Conservative 56; Mismatches 177; Indels 88; Gaps 19;

```
QY 3 VAR---GIYKADVAQS-----SYGLYGGQIVAVADTGLDTGRNDSMHEAFRG 48
Db 160 VARVWLDGVKASLDTSVGQIGTEPKAWEAGYDGGVXIALDTGV-----ATHPDLKG 213
QY 49 KITALLALGRNNANDPNEGHTVHAGSVLNGA---TNKGMAPQANLVFQSIMDSSGGL 104
Db 214 QVTASKNFTSAPPTGGDVVGHGTHVASTAAGTGAQSKGYKGVAPGAKILNGKVLDDAG-- 271
QY 105 GGLPSNLQTLFSSQFSAGARIHTNSWGAAVNGAVTTDSRVVDYRK--NDMTILF--AA 160
Db 272 FGDDSGILAGMENAAGADIYVMSLG---GMDTPETDPLEAAVDKLSAEKILFATA 327
QY 161 GNEGPNGTISAPGTAKNATTVGATENLRPSFGSYADNINHVAFSSRGP-TKDGRIKPD 219
Db 328 GNEGPQ--SIGSPGSADSLTVGA-----VDDKDKLADFSTGPRLDGAVKPD 374
QY 220 VMAPGTIILARSGLAPDSSEFMANHDSKYAVMGTSNAPPIVAGNVAVOLREHFVKRGIT 279
Db 375 LTAGVDTITAAAKGNDIAKEVEKEPAGYMTISGTSNATPHVAGAAALLKQCHPE----- 429
QY 280 PKPSLLKALLIAGADVGLGYENGNGWGRVTLDKSLNAVYNSSALS----- 328
Db 430 WKYAEELKAGALTASTKDQ--KYTFPEQSGRVDYKAITQVIAEPVLSFVGVOQWPHADD 487
QY 329 --TSOKATYFTETAGKPKLISLWSD-----APAS--TTASVTLVNDLDVITAP-NGT 378
Db 488 KPYTKLITYRNLGTEDVTLKLTSTAIGPKKAAAPAGEFTLGASTL-----TVPANGT 539
QY 379 RYVGNDPSAPFDNNMDGRNNVENVFINSPOS---GTYTIEVQANYV 421
Db 540 ASYDVVTADTRLGAVDGTYSAYVVAATGAGQSVRTAAVEREVESYV 586
```

Search completed: March 31, 2004, 16:33:58
Job time : 41.2177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:38 ; Search time 11.3508 Seconds
(without alignments)
(3677.911 Million cell updates/sec)

Title: US-09-985-689A-2
Perfect score: 2245
Sequence: 1 NDVARGIVADVAQSSYGLY.....EYQAVNVDPGPNLSLATVN 434

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_78:***
1: p1r1:***
2: p1r2:***
3: p1r3:***
4: p1r4:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	526.5	23.5	1743	2 T18279	multidrug resist
2	490	21.8	1905	2 T18267	multidrug resist
3	357.5	15.9	444	2 B83891	intracellular alka
4	326	14.5	442	2 A69587	intracellular alka
5	302.5	13.5	806	2 A41341	microbial serine p
6	302.5	13.5	1398	2 T28159	pyrolysins (EC 3.4.
7	283	12.6	419	1 S28535	subtilisin (EC 3.4
8	279	12.4	799	1 G83753	subtilisin-type pr
9	278	12.3	420	1 S23407	subtilisin (EC 3.4
10	277	12.3	580	2 S11890	serine proteinase
11	274.5	12.2	1345	2 T29090	surface layer-asso
12	270	12.0	715	2 JG4808	alkaline serine pr
13	260	11.6	1331	2 A72547	probable surface l
14	259.5	11.6	513	1 A35742	equalysin (EC 3.4.
15	259	11.5	627	1 D75393	serine proteinase,
16	254.5	11.3	757	2 C84120	subtilisin-type pr
17	252.5	11.2	534	1 J50173	alkaline proteinas
18	250.5	11.2	402	1 JU0330	alkaline proteinas
19	249.5	11.1	894	2 P69730	cell wall-associat
20	245	11.0	380	2 A49778	high-alkaline seri
21	245.5	10.9	401	2 T39874	serine proteinase
22	241	10.7	488	2 A11930	proteinase (import
23	240	10.6	378	2 A33973	high-alkaline seri
24	239	10.6	382	2 I39780	subtilisin (EC 3.4
25	238	10.6	519	2 S71451	halolysin R4 (EC 3
26	234.5	10.4	379	1 SUBSCL	subtilisin (EC 3.4
27	234.5	10.3	382	1 SUBSN	subtilisin (EC 3.4
28	232	10.3	321	1 S27501	alkaline proteinas
29	231.5	10.3	1052	2 T17093	intraluminal subti

30	230.5	10.3	613	2 S75976	hypothetical prote
31	229.5	10.2	381	2 JH0778	subtilisin (EC 3.4
32	229	10.2	1167	1 A35066	streptococcal C5a
33	228.5	10.2	525	2 G84406	halolysin (importe
34	228	10.2	1374	2 D72593	hypothetical prote
35	227.5	10.1	381	1 SUBS81	subtilisin (EC 3.4
36	226	10.1	384	2 JG4802	alkaline proteinas
37	225.5	10.0	272	2 A23624	subtilisin (EC 3.4
38	225.5	10.0	275	2 JG1085	subtilisin (EC 3.4
39	225.5	10.0	381	1 SUBS8	subtilisin (EC 3.4
40	225.5	10.0	381	2 JQ1487	subtilisin (EC 3.4
41	225.5	10.0	910	2 C69456	subtilisin sendai
42	224	10.0	535	2 B82358	alkaline serine pr
43	222.5	9.9	769	2 D86335	T20H2 6 protein
44	219.5	9.8	274	1 SUBSD	subtilisin (EC 3.4
45	219.5	9.8	361	2 A48373	high-alkaline seri

ALIGNMENTS

RESULT 1
T18279 multidrug resistance transport protein - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18279
R:Shaulsky, G.; Loomis, W.F.
Submitted to the EMBL Data Library, June 1996
A:Reference number: Z18855
A:Accession: T18279
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1743 <SHA>
A:Cross-references: EMBL:U60086; NID:G1399914; PID:G1399915; PIDN:AA03331.1
C:Genetics:
A:Gene: tagC

Query Match 23.5%; Score 526.5; DB 2; Length 1743;
Best Local Similarity 27.8%; Pred. No. 9.8e-26;
Matches 164; Conservative 82; Mismatches 152; Indels 191; Gaps 22;

QY	19	LYGGQIVAVADTGLDGR---NDS-----SMHEFRKRTALVALGRTNNANTNGH	68
DB	314	LRGGQILSLADTGLDSCFCFSDSKYPIPLNSVNLNHR-KVVTYTTSTSDSDKVDGH	372
QY	69	GTHVAGSVLG-----NGATNKGMAPQANLVFQSIWDSGGGHPNLTQTFQSAFSA	121
DB	373	GTHICGSAAGTPEDSSVNISSFGLATDAKIAFPDLASGSSSLTP--PSDLKOLYQPLYDA	431
QY	122	GARIHNSWGA---AVNGAYTTDSRVVDYERKN-DMTLLFPAAGNBPVGGTIS--APG	174
DB	432	GARVHCDMSGVSVEGYTGYSSTASIDPLFTHPDPFILRAAGN---NEQYLSILTQS	488
QY	175	TAKNAITVGAATENR-----PFGSYADNI-----	199
DB	489	TAKVITVGAHQTHENVYLDGRVNTYQSSVDINQSLIDPFSRYCNVTTAQCLESNA	548
QY	200	-----NHVAQFSSRGGTQKGRIPEDVMAQPTYLISA	230
DB	549	TTGLASCCPTLLRKSVIDAANTOPLLYNENNISFSSKGPETHGRMPALVAGGEYITSA	608
QY	231	RSSLA-----PDSSFWANHSKAYVAGTSMATPIVAGVAOLREH-----FVK	274
DB	609	RSNANNTDQCGDSL--PNNNALIA--ISGISMATSPAAATTLIKQYLVGVGYPTGSIVE	666
QY	275	NRGITPKRSILKAALIGA-----ADVGLGYPNGN-----QGWGRVTL	313
DB	667	SNKQPTGSLIKALMINNAOLNGLTQTLTSSITVPSNVFENFAGASLVQGMGAIKMS	726
QY	314	KSLVAVVNESS-----ALTSQAKATYTF-----	338
DB	727	NMLHVAVNNNSNNNNKTSQGITVFDGIGGLDLVKNQMKESLSTGQNTSYCFITYKPS	786

QY 339 ---ATACKPKL---ISLWSDAPASTTASVTLVNDLDTLTAPNGRTYVGNDFSAFDDNN 392
 Db 787 SSSSSSSGNIPRVATLTWTDPPSYAGAKFNLVNNDLTLTI-----YRDNOSTIFYEN 840
 QY 393 WDGR-----NNVENVFINSPOSGTYTIEVQANVPVGPONFS 429
 Db 841 QGSSSFLGLAPTQDTLANNBGIYVNPTEPWTTFPMVAGIVNPVGPONFS 889

RESULT 2

multidrug resistance protein - slime mold (Dictyostellium discoideum)
 C:Species: Dictyostellium discoideum
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T18267
 R:Shaulsky, G.; Kuspa, A.; Loomis, W.F.
 Submitted to the EMBL Data Library, January 1995
 A:Description: An MDR transporter/serine protease gene is required for prestalk special
 A:Reference number: Z18850
 A:Accession: T18267
 A:Status: preliminary; translated from GS/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-1905 <SHA>
 A:Cross-references: EMBL:U20432; NID:G664839; PID:G664840; PIDN:AAA62212.1
 C:Genetics:
 A:Gene: tagB

Query Match 21.8%; Score 490; DB 2; Length 1905;
 Best Local Similarity 28.0%; Pred. No. 2,4e-23;
 Matches 162; Conservative 74; Mismatches 159; Indels 184; Gaps 23;

QY 19 LYGGQIVAAVADTGLDTCR---NDS-----SMHEAFRGKITALYALGRTNANDTNGH 69
 Db 376 LRGGQILSTADTGLDSDSHCFPSDKYRPFQVQNEHRAVYV---YITHDNEDEVNQH 432
 QY 69 GTHVAGSVLNG-----ATNKMAPQANLVFQSIHDSGGIGL-PSYLQTLFEOQAS 120
 Db 433 GTHVCGSAAGPEDSSVAISSFGLATDAKIAFYDL---SSGSSEPTPEDEYQMYPLVD 490
 QY 121 AGARHTNSMGA---AVNGAYTTDSRVNDVYRK-NDMTILFAAGNEGPGGTISAPGT 175
 Db 491 AGARVHDSDWGSVGLQGYGYGSDDAGIDAFLEYEFESILLAAQN-NELFASLSLAQNT 549
 QY 176 AKNAITGATENLSPFGS-----YADNI----- 199
 Db 550 AKNAITGAEQTAHVNVYSDALEYDPSDNANFQPCLPDKKYCNYYTAKCCEVSNVKG 609
 QY 200 -----NHVAFSSRPCTKDGRIKPDVMAPGTYILSARS 233
 Db 610 LQCCPASIKQNASDFTTQPOPYNNMNGSFSSKGTTHGRILKPDIVAGEYITTSARN 669
 QY 234 -----IAPSSFWANHDSKRYAVGTSMAPIVAGNVALRHF-----YKNRG 277
 Db 670 GENSTDOCQGGSL--PVANGILMSISGTSMAPIPLATATTLRQYLVGVPTGESVEENK 727
 QY 278 ITPKPSLLKALILGADVGLGY-----PNGQMGRTVLDSLVN 318
 Db 728 LIPGSLILKALMINNQLNGTYWSSSTNPSNAITEQINGANLQGMALNN--NW 784
 QY 319 AYVNESS-----ALTSOKAT-----YFFT-----ATAG 342
 Db 785 LYVKSNSPTPSRWIGIGLQKQKATEWKEDLSGLNKSYCFYKPSSSSGSGGGG 844
 QY 343 KP-LKISLVMSDAPASTTASVTLVNDL-----VITAN--GTRYVGNDFSAFDD 390
 Db 845 TPRVATLTWTDPPSYGAKFNLVNNDLILNSDDSLITIGNGSGSLQPAKVAQP-- 902
 QY 391 NMWDGRNVNVPINSPOSGTYTIEVQANVPVGPONFS 429
 Db 903 ---DTLNNVEGIIINPTKANMYKFTIAGTVNPVGPONFS 938

RESULT 3

B83891
 Intracellular alkaline serine proteinase aprx [imported] - Bacillus halodurans (strain
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C:Accession: B83891
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir
 Nucleic Acids Res. 28, 4317-4331, 2000
 A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A:Reference number: A83650; MUID:20512582; PMID:11058132
 A:Accession: B83891
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-444 <STO>
 A:Cross-references: GB:AP00513; GB:BA000004; NID:G10174345; PIDN:BA05649.1; GSPDB:GNO
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: aprx

Query Match 15.9%; Score 357.5; DB 2; Length 444;
 Best Local Similarity 30.2%; Pred. No. 1.2e-15;
 Matches 110; Conservative 52; Mismatches 105; Indels 97; Gaps 16;

QY 11 DVAQSYGLYGGQIVAAVADTGLDTCGRNDSMHEAFRGKITALY-ALGRTNANDTNGH 69
 Db 135 EVIRNGERTLTGKDVTIANVIDTGI-----YPHDELGRIKAFVDVFNQREBPYDNGHG 187
 QY 70 THVAGSVLNGATN-----KMAPQANLVFQSIHDSGGIGLPSNLO----- 112
 Db 188 THCGADAAAGASSDGGYRGFPAPFANVIGYVLNKKQ-GMGSLESIMQVEMCTIQNEHP 246
 QY 113 -----TLFSQAF-----SAGARHTNSMGAALVNGAYTTDSRVNDVYRKNDMTI 156
 Db 247 DDPHILISMSLGGQALPVENQEDPMVRIVEBANNAIG-----TV 286
 QY 157 LPAAGNEPNGGTISAGTAAATTTGATENTLPSFSYADNINHVAGFSSRGPTKGR 216
 Db 287 CVAAGNSGPDQAQTIASPSVSEKVTIVGALDD-RITRREDD--VAFSSRGPTTYGKP 342
 QY 217 KPDVAPGTYYLSARSSILAPDSF-----WANHDSKRYAVGTSMAPIVAGNVALRHF 271
 Db 343 KPDILAPGVNVLSLR---PVSFYDKIQKSGRVSHNTMNSGTSMAPIVAGNVALRHF 399
 QY 272 FVKRNGITPKPSLLKALILGADVGLGYPNGQMGRTVLDSLVN--AYVNESSALS 328
 Db 400 ---EPNLTPDE--VKTRLM-----EGTDRWA---DRDPNYGAGYISABGAIP 439
 QY 329 TSOK 332
 Db 440 NSEE 443

RESULT 4

A69587
 Intracellular alkaline serine proteinase aprx - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
 C:Accession: A69587
 R:Kumari, F.; Ogata, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
 C:Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A:Enrich, S.D.; Emerson, P.T.; Enliat, K.D.; Errington, J.; Fabret, C.; Ferrati, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
 ich, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapides, A.; Lardinois
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.W.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y, M.; Ogawa, K.; Ogata, A.; Oudega, B.; Park, S.H.; Patro, V.; Pohl, T.M.; Portetelli
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
 akuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
 A:Authors: Yoshikawa, H.F.; Zundstein, E.; Yoshikawa, H.; Zundstein, A.
 A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: A69587
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-442 <KUN>
 A:Cross-references: GB:299113; GB:AL009126; NID:92634090; PIDN:CAB13610.1; PID:e1183385;
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: aprX
 C:Superfamily: subtilisin homology
 F:146-398/Domain: subtilisin homology <SBT>
 Query Match 14.5%; Score 326; DB 2; Length 442;
 Best Local Similarity 30.6%; Pred. No. 1,2e-13;
 Matches 103; Conservative 45; Mismatches 97; Indels 92; Gaps 14;
 QY 11 DVAQSSYGLYGGGQYVAADTGLDGRDSSMHEAFRGKITALVALGRTNNAN-----D 64
 Db 136 EVVRNGQTLTGKGVTAAVDTGI-----YHPDLERI-----IGPADVWQKTEPYD 183
 QY 65 TNGHGTVAAGSVLGGGATN---KGMAPQNLVFGSINDSSGGLGLPSNLTQLFSQAFS 120
 Db 164 DNGHGTGAGDVASSGASSGGYRGPAPBANIIGVKVANKGS-CLTLDILEGV----- 236
 QY 121 AGARLHTNSWGAAN-----GAYTTSKRVVDYVRKND-----MTIL 157
 Db 237 -----EMCTQYEDNDPEPIDMSMSLGGDALRYHEGDPVRAVEASAGIYVC 288
 QY 158 PAAGNEGPGGTISAPGTAKNAITVATENLEPFSGVADINHTVACSSSGPTKDGRIK 217
 Db 289 VAAAGNSGPGSQTIAAPGUSEKVTITGALDNNMTA-----SSDDTYASSSSSGPTIYGEK 344
 QY 218 PDVMAPGTYILASRSLAPDSGF---WANHSRYAVMGTSWATPIVAGNVAGLREHF 272
 Db 345 PDILAPGVNIIISLR---PNSYIDRLQKSSRVSQYFTMSGTSMATPICAGIALIILQ-- 399
 QY 273 VKNRGITPK--PSILKALINAGADVGLCPNGNGM 307
 Db 400 -QNPDLTPDEVKELK-----NGTTKW 420
 RESULT 5
 A41341
 Microbial serine proteinase (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 17-Jul-1992 #sequence revision 17-Jul-1992 #text change 20-Jun-2000
 C:Accession: A41341; B41341; S39700; D69730
 C:Residues: 1-173; 6889-6895, 1991
 J:Rioma, A.; Ruto Jr., G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Pero, J.
 A:Title: Cloning and characterization of the gene for an additional extracellular serine
 A:Reference number: A41341; MUID:92041574; PMID:1938892
 A:Accession: A41341
 A:Molecule type: DNA
 A:Residues: 1-806 <SLO>
 A:Cross-references: GB:M76590; NID:9143819; PIDN:AAA2881.1; PID:9143820
 A:Accession: B41341
 A:Molecule type: protein
 A:Residues: 161-195 <SL2>
 R:Glaeser, P.; Kunst, F.; Arnaut, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu,
 A.; Rapoport, G.; Darchin, A.
 Mol. Microbiol. 10, 371-384, 1993
 A:Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region in
 A:Reference number: S39655; MUID:95020537; PMID:7934828
 A:Accession: S39700
 A:Molecule type: DNA
 A:Residues: 1-806 <GLA>
 A:Cross-references: EMBL:X73124; NID:9413933; PIDN:CAA51601.1; PID:9580871
 R:Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Bruchet, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
 A.; Enllich, S.D.; Emmerson, P.T.; Entian, K.D.; Erington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
 Tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y., M.; Ogata, K.; Ogatawara, A.; Oudega, B.; Park, S.H.; Paro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
 A:Authors: Schleicher, S.; Schreier, R.; Scifone, F.; Sekiguchi, J.; Sekowska, A.; Serot
 Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, i
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Darchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: D69730
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-806 <KUN>
 A:Cross-references: GB:299123; GB:AL009126; NID:92634240; PIDN:CAB15835.1; PID:9263444
 A:Experimental source: strain 168
 C:Comment: The amino terminal sequence of the mature protein and a molecular weight of
 C:Genetics:
 A:Gene: vpr
 A:Start codon: TTG
 C:Superfamily: microbial serine proteinase vpr, subtilisin homology
 C:Keywords: hydrolase; serine proteinase
 F:1-28/Domain: signal sequence #status predicted <Sig>
 F:29-160/Domain: propeptide #status predicted <Pro>
 F:160-348/Domain: subtilisin homology #status atypical <SBT>
 Query Match 13.5%; Score 302.5; DB 2; Length 806;
 Best Local Similarity 23.2%; Pred. No. 8.7e-12;
 Matches 127; Conservative 58; Mismatches 161; Indels 201; Gaps 17;
 QY 18 GLYGQGIYVAADTGLDTR-----NDSMHEAFRGKITALVALGRIN 60
 Db 177 GYTGKGIKVAIDTGVVHNPDIKKNFGQYKGYDVNDYDPERPTG-----D 225
 QY 61 NAMDTHGHTVAAGSVLGGATNKGAPQNLVFGSINDSSGGLGLPSNLTQLFSQAFS 120
 Db 226 PRGENDHGTVAAGTAAAG-TIKGVAADTLLAVLVLPGG--SSTENVIAGVERAVQ 282
 QY 121 AGARLHTNSWGAANVATDTSKRVNDYVRKNDMTILFAAGNEGPGGTISAPGTAKNAI 180
 Db 283 DGADVNNLSLGNLNPDMATSTAL-DWMSSEGVAVTNSNGSPNGWVGSPGTSREAI 341
 QY 181 TVGATE---NLSPFGSY----- 195
 Db 342 SVGATQPLNEVAVTGTSGSYSAKVMGYNKEDDYKALNKEVELVAGIGEAQDFEGKDLT 401
 QY 196 -----ADNI----- 199
 Db 402 GKVAIVREGSIAFVDKADNKKAGAGIMVYNNLSGEIEANVPQMSVPTIKSLDEGKLT 461
 QY 200 -----NHVAFSSRGPTK-GRIKPVMAPGTYILASRSLAP 236
 Db 462 VSALKAGETTTKTLVSKALGQVADPFSRGVDMTMIKPIASPGNNIVSTIITHP 521
 QY 237 DSEFWANHSRYAVMGTSWATPIVAGNVAGLREHFVKNRGITPKSL-LKALTAGAA 294
 Db 522 D-----HPYGSQKQSTWASPHIAGAVIRK-----AKPKMSVEQIKALINMTAV 568
 QY 295 DV-----GLGPNQNGGRTLLDKSLNVAVNNSALSTSQKATYFTTAGAPLISLV 350
 Db 569 TLKQSDSEVTPHNAQGSARI---KNA--IKADSLVSPGSIYGFILKNGMETKNETF 623
 QY 351 WSDAPASTASTVTLVNDLVLITAPNGTRVYVNDFAFPDNNMDGNNVENVFINSPOG 410
 Db 624 TIENQSSIRKSYTL-----EVS--FNGSGISTSGTSRAVVI PAHOTG 662
 QY 411 TYTTEVQ 417
 Db 663 KATAKVR 669
 RESULT 6
 T28159

pyrolysin (EC 3.4.21.-) - *Pyrococcus furiosus*
 C:Species: *Pyrococcus furiosus*
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 24-Oct-2000
 C/Accession: J28159
 R/Voucher: W.G.B., Eggen, R.I.L., Geerling, A.C.M., Platteau, C., Sieren, R.J., Vos, J. Biol. Chem. 271, 20426-20431, 1996
 A>Title: Isolation and characterization of the hyperthermostable serine protease, pyrolysin
 A/Reference number: J28159
 A/Accession: J28159
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1398 <VOO>
 A/Cross-references: EMBL:U55835; NID:G1556462; PID:G1556463; PIDN:AA809761.1
 A/Experimental source: DSM3638
 C/Genetics:
 A/Genes: ple
 C/Keywords: hydrolase; serine proteinase

Query Match 13.5%; Score 302.5; DB 2; Length 1398;
 Best Local Similarity 26.3%; Pred. No. 1.8e-11;
 Matches 138; Conservative 57; Mismatches 177; Indels 153; Gaps 22;

21 GGGIVAVADTGTGTRGDS-----SMHEAFRGKITALALGRTNNAN-----63
 Db GNGVDLYVDITDIDYDTDEVPGLQVNTVDVAVFSYVYGLVLAIEDPGEYAVFGW 360
 QY 64 DTNGGTHVAGSVLGNGATN-----KGMAPQAN 91
 Db DGHGHTHVAGTVAGVYDSNNDAMDWLSMYSGEWEVESRLYGMQDYNTVTTDTVGVAFGA 420
 QY 361 DGHGHTHVAGTVAGVYDSNNDAMDWLSMYSGEWEVESRLYGMQDYNTVTTDTVGVAFGA 420
 Db 92 LVPQSVDSGGGLGSPNSIQTLFSAFSGARITHNSMAANGAYT--TDSRN--VDD 147
 QY 421 IMAIRVL-RSDGRSMWDITIEGM-TYATHGADVISMSLGG--NADYLDGTPESVAVDE 476
 Db 148 YVRKNDITILFAAGNEGPNGTISAPGTAKAATVGTATENTLRPSFGSYAD-----197
 QY 477 LTRKYGAVFVLAAGNEPGINIVSGPVATKATITVGAA--VPINGVYVYQALGPDYVG 535
 Db 198 -----NINHYAQFSSRGPTKDGRIKPDVMAFGTYILSARSLAPDSSFPANHDSKYA 249
 QY 536 FYVFPAYTNV-RIAFSSRGPRIDGELKPVVAFGVYSSLSLPMWIGADF-----585
 Db 250 YMGSTMATPIVAGNVAAQLREHFVKNGITPKPSLKAAALAGAA-----DVGIGYPNG 303
 QY 586 -MGSTMATPHVSGVAVALLISG--PKRPGIYVNDIILKYLESGATVLEGGPYGQCYTEL 643
 Db 304 NQCKGRVTLDKSLNAVAVNESSALSTSQKATYFTTAGKPKLSLWSDAPASTTA---360
 QY 644 DQGHGVNVTKSWEI-----LKAINGTTLPIVDHMADKSYSDPAEYL 685
 Db 361 SVTLVNDLDTVITAPN---GTRYVGNDFSAFPD---NNW-----DG-----RNNVE---400
 QY 686 GVDVIRLVLNANSPDILVEWHIKYVGDTEKTRFEIYATEWIKRPFVSGVILENTEFVL 745
 QY 401 -----NVFTNSPQSGTYTIEVQAVNVPNGPQNF 429
 Db 746 RVKVDVGLPGLVYGRRIIDDP--TPVIEDEILNTIVPEKFT 788

RESULT 7
 S25835
 subtilisin (EC 3.4.21.62) precursor - *Bacillus* sp. (strain TA41)
 C/Species: *Bacillus* sp.
 C/Date: 22-Nov-1993 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999
 C/Accession: S25835
 R/Davall, S.; Feller, G.; Narinx, E.; Gerday, C.
 Gene 119, 143-144, 1992
 A>Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph *Bacillus*
 A/Reference number: S25835; MUID:93012986; PMID:11398082
 A/Accession: S25835
 A/Molecule type: DNA
 A/Residues: 1-419 <DVV>
 A/Cross-references: EMBL:X63533; NID:G40198; PIDN:CAA45096.1; PID:G40199

C:Superfamily: subtilisin; subtilisin homology
 C/Keywords: extracellular protein; hydrolase; serine proteinase
 F1-23/Domain: signal sequence #status predicted <SIG>
 F124-110/Domain: propeptide #status predicted <PRO>
 F111-419/Product: microbial serine proteinase #status predicted <MAT>
 F135-373/Domain: subtilisin homology <SBT>
 F144,184,359/Active site: Asp, His, Ser #status predicted

Query Match 12.6%; Score 283; DB 1; Length 419;
 Best Local Similarity 33.0%; Pred. No. 6.6e-11;
 Matches 87; Conservative 37; Mismatches 98; Indels 42; Gaps 12;

21 GGGIVAVADTGTGTRGDSMH-EAFRGKITALYALGRT--NNADTNGHTHVAGSV 76
 Db GAGINIAVLDGVNTNHPDLSNNVEQCKD-----FTVGTFTNSCTDRGQGHVAGSA 189
 QY 135 GAGINIAVLDGVNTNHPDLSNNVEQCKD-----FTVGTFTNSCTDRGQGHVAGSA 189
 Db 77 LGNATNK---GMAPQANVYQSIM--DSSGGLGSPNSIQTLFSAFSGARITHN--SW 130
 Db 190 LANGTSGVYGVAPEDLWAAVKVLGDDSGGYADDAEALRHAGDQATLNTVYIMSL 249
 QY 131 GAAYNCAVTTDSRVVDYVERKNDITILFAAGNEGPNGTISAPGTAKAATVGTATENTLRP 190
 Db 250 GSSGESSLIT---NAVYAVDKGVLITAAAGNSGPKRSGISGYGALVNVAVALLENTIQ 306
 QY 191 SFGSYADNINHYAQFSSRGPTKDG-----RIKPDVMAFGTYILSARSLAPDSSFPANH 244
 Db 307 N-GTY-----RVAQFSSRGKRRTAGDYVIQGDVEISAPGAAYST-----W--F 348
 QY 245 DSKYAYMGSTMATPIVAGNVAAQL 268
 Db 349 DGYATISGTSMASSPHAAGLAKI 372

RESULT 8
 G83753
 subtilisin-type proteinase (EC 3.4.21.-) vpr precursor [similarity] - *Bacillus halodura*
 C:Species: *Bacillus halodura*
 C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C/Accession: G83753
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saeaki, R.; Masui, N.; Fuji, F.; Hir,
 Nucleic Acids Res. 28, 4317-4331, 2000
 A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A/Reference number: A83650; MUID:20512582; PMID:11058132
 A/Accession: G83753
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-799 <STO>
 A/Cross-references: GB:AF001510; GB:BA000004; NID:G10173440; PIDN:BA804550.1; GSPDB:GNO
 A/Experimental source: strain C-125
 C/Genetics:
 A/Genes: vpr
 C:Superfamily: microbial serine proteinase vpr; subtilisin homology
 C/Keywords: hydrolase; serine proteinase
 F1-29/Domain: signal sequence #status predicted <SIG>

Query Match 12.4%; Score 279; DB 2; Length 799;
 Best Local Similarity 25.0%; Pred. No. 2.8e-10;
 Matches 119; Conservative 47; Mismatches 134; Indels 176; Gaps 17;

18 GLVGGQIVAAVDGTLTGTRGDSMHAFRGKITALALGRTNNANDT-----NGH 68
 Db 171 GYTGSGITVALDGVDTHTPD--LVHAF--GDYKGMDFINNDPQRTPPGDPRIETTH 227
 QY 69 GTHVAGTVANGATNKGMAPQANLVFQSIDSSGGLGSPNSIQTL--FSAFSGARITH 126
 Db 228 GTHVAGTVANGATNKGMAPQANLVFQSIDSSGGLGSPNSIQTL--FSAFSGARITH 126
 QY 127 TNSMAANGVYTTDSRVVDYVERKNDITILFAAGNEGPNGTISAPGTAKAATVGTATE 186
 Db 283 NISLGNITNDPDPFATSTAL--DMAAEGVAVTSSNGSGPNNWTVGSPGTSRDAISVAT- 340
 QY 187 NLKPSFGSY-----ADNINH-----201

Db 341 --RLPNKYKAVFTSDGIDYPSADJINGPSPDELLDGEYEVAFAGLKPDEGEVD 398
QY 202 ----- 201
Db 399 VEGKIALIVAGEIPFEKAKNAKAAGVGAIIYNNVAGVPTVGLAIPITMISNDGLK 458
QY 202 -----VAQSSRGPT-KDGIKPDVAPGTYILSARSSLA 235
Db 459 MRNELENGQNTVTFIEPDXLVEGETVADFPSSRGVMTWIKPDVAPGVAIVSTIPTHQ 518
QY 236 PDSEFMANHDSKYAVMGSTMAPPIVAGNAQLREHVNRKRGITPRSLKALLINGAD 295
Db 519 PDDPY-----GYSRGITSMASHRAGAAALLIEAH-FWVGV---DHVKALLMTAEN 567
QY 296 V---GLGYPNGQNGWGRVTLDRSLNVAAYNESSALSTQKATY-TFTATAGKPLK 346
Db 568 LVDENGRNPYPHNTQAG-----SRIYDAIESETLVTGSHGFGFTTKERGRQVE 617

RESULT 9

Subtilisin (EC 3.4.21.62) 1 precursor - Bacillus sp. (strain TA39)
C:Species: Bacillus sp.
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 18-Jun-1999
C:Accession: S23407
R:Marink, B.; Davail, S.; Feller, G.; Gerday, C.
Biochim. Biophys. Acta 1131, 111-113, 1992
A:Title: Nucleotide and derived amino acid sequence of the subtilisin from the antarctic
A:Reference number: S23407; MUID:92256481; PMID:1581352
A:Accession: S23407
A:Molecule type: DNA
A:Residues: 1-420 <NR>
A:Cross-references: EMBL:X62363; NID:940200; PIDD:CAA44227.1; PTD:940201
C:Gene: stud1
C:Superfamily: subtilisin; subtilisin homology
C:Keywords: extracellular protein; hydrolase; serine proteinase
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-111/Domain: signal sequence #status predicted <PRO>
F:112-420/Product: microbial serine proteinase #status predicted <MAT>
F:116-374/Domain: subtilisin homology <SPT>
F:145,185,360/Active site: Asp, His, Ser #status predicted

Query Match 12.4%; Score 278; DB 1; Length 420;
Best Local Similarity 33.2%; Pred. No. 1,4e-10;
Matches 88; Conservative 38; Mismatches 95; Indels 44; Gaps 13;

QY 21 GGGQIVAVADTGLDTGRNDSMEAFRGKITAL-VALGT--NNANDTNGGHTVAGS 75
Db 136 GGGGINIVLDGTGVTN-----HDLNNVEQCKDFVGTYYNNSTCTROGHGTIVAGS 189
QY 76 VLNGATNKK--GNAPQNLVFOGIM--DSGGIGGLPSNLQTLFSGAFAGARIRHTN-S 129
Db 190 ALADGGGNGVYGVAPADLMAYKVLGDGSGVADDAIRARAGQATLNTKVTNMS 249
QY 130 WGAIVNAGVYTTDSNVDVYKNDMTLLPAAGBENGNGTISAPGAKNAITVGAENIR 189
Db 250 LGSSGESLITNAVN--YNNKQVLIIAAGNSGPYQSIGYPGALVNAVAALLEN-K 305
QY 190 PSFSGSYADNINHVAFSSRGPT-KDG-----RIKPDVMAAGTYILSARSLADSSFMAN 243
Db 306 VENTY-----RVADFSRGYSWTDGYAIQKQDVEISAPGAALYST-----W-- 348
QY 244 HDSKYAVMGSTMAPPIVAGNAQL 268
Db 349 FDGGYATISGTSMAVSPHAGLAAKI 373

RESULT 10

S11890
serine proteinase (EC 3.4.21.-) precursor, extracellular - Xanthomonas campestris pv. ca
N:Alternate names: subtilisin-related proteinase
C:Species: Xanthomonas campestris pv. campestris

C:Date: 21-Nov-1993 #sequence_revision 07-Feb-1997 #text_change 03-Dec-1999
C:Accession: S11890
R:Lin, Y.N.; Tang, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
Mol. Gen. Genet. 220, 433-440, 1990
A:Title: A multipurpose broad host range cloning vector and its use to characterise an
A:Reference number: S11890; MUID:90251253; PMID:2187155
A:Accession: S11890
A:Molecule type: DNA
A:Residues: 1-580 <LID>
A:Cross-references: EMBL:X51635; NID:948533; PIDD:CAA35962.1; PTD:948534
A:Experimental source: Xanthomonas campestris pv. campestris
A>Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-AL
C:Superfamily: subtilisin; subtilisin homology
C:Keywords: extracellular protein; hydrolase; serine proteinase
F:1-32/Domain: signal sequence #status predicted <SIG>
F:168-423/Domain: subtilisin homology <SPT>

Query Match 12.3%; Score 277; DB 2; Length 580;
Best Local Similarity 26.6%; Pred. No. 2.5e-10;
Matches 126; Conservative 53; Mismatches 150; Indels 144; Gaps 22;

QY 21 GGGQIVAVADTGL-----DTGRNDSMEAFRGKITALVALGRTYNNAND----- 64
Db 166 GGGIVAVADTGLTSHDNLNNILAGDFISDATTAADQNGRDNADDEGDMVAANCGA 227
QY 65 -----TNGGHTVAGSVLNGCATNKGMAFOA-----NLVFQSTMS 100
Db 228 GIPAASSWGHGTHVAGVAAVNTNTGVAGVAGKVPVAVLCKGGSLSIDADIYMA 287
QY 101 SGG-UGGAPSNLQ--TLPSQAFSAGARHTNMGAAVAGVATTSRVVDVYRKNDMTIL 157
Db 288 SGGTYSGLPAAVAPRAVYVNSLGGGSGSTTMQN-ALNGAISRST-----TVV 334
QY 156 FAAGNEGNGGTTAPGTAKAIVGATEN--LEPSGSYADNINHVAFSSRGPTDGR 215
Db 335 VAAQNDASNVGQ-SIPANCANVIAAATTSAGAAVSXNFGTGL----- 377
QY 216 IKPDVMAAGTYILSARSS--LAPDSFMANDSXYAVMGSTMAPPIVAGNAQLREHFV 273
Db 378 ---DVSAGSILLSTLNSGTTTPOS-----ASYASNGTSMAVHAGVVALVQS--V 425
QY 274 KNRGTPPK--PSLKK--AALTAGADVGLG-----PENGQGWRTLTK 314
Db 426 APTALTTPAAVETLLKNTARLPACSGCGGAGVNAADAATAALNGSGGGGGGNTLTN 485
QY 315 SLNVAAYNESSALSTQKATYFTATAGKPKIKSLVMSDAPASTASVTL--VNDLDIV 371
Db 466 GTPVT---GLGATGALNLTITVPAG-----SGTLVTTISGSGGDADLY 527
QY 372 ITAPNGTRYVNDPSAPFDNNMDR---NNVENVFINSPOSGTYTTEVOAYN 420
Db 528 VRAG-----SAPTDSAYTCRPYRSGNAETCTITAP-SGYVYRLKAYS 569

RESULT 11

T23090
Nucleic acid-associated STABLE proteinase - Staphylococcus marinus
N:Alternate names: hyperthermostable proteinase
C:Species: Staphylococcus marinus
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T23090
R:May, J.; Lupas, A.; Kellermann, J.; Eckerskorn, C.; Baumeister, W.; Peters, J.
Curr. Biol. 6, 739-749, 1996
A:Title: A hyperthermostable proteinase of the subtilisin family bound to the surface lay
A:Reference number: 220559; MUID:96385442; PMID:8793300
A:Accession: T23090
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1345 <MAY>
A:Cross-references: EMBL:U57968; NID:91374755; PTD:91374756; PIDD:AA032323.1
A:Experimental source: strain F1
C:Function:
A:Description: probably serves an exodigestive function related to the organism's energy

A>Note: stoichiometric S-layer component

Query Match 12.2%; Score 274.5; DB 2; Length 1345;

Best Local Similarity 29.3%; Pred. No. 1.1e-09;

Matches 97; Conservative 48; Mismatches 105; Indels 81; Gaps 14;

46 FRKITALYLGRTNNANDTNGHCTHA-----GSLV-----GNGKTNK--GMAQANLV 93

445 YQGRYIAL-----VSDFHGHTSVATVIASRGVLDYDGKLYRIMGAPAKI- 495

94 FQSIMDSGGLGGLPSTLQTLFQAFSAG-----ARHTNSM 130

496 -----AGSDAMLLGNILVL--EAWLAGENITEBEDGVYILSDPFPFHADIISNM 546

131 GAANVNAVITDSSKND-----DYAKNDMTLLEAAGNENGGITTSAPGAK 177

547 GSIIYINFWLQFPGLIRSSFMDEILAIRNYLLGDHVTIVFAAGNGPFGYSNGAPGTL 606

178 NATTVGATE--NLRPSFG--SYADNINVAQPSRSGPTKGRIRKPDVAPGTIILSARS 232

607 LVITAGASTLMDYTRIGYPEGAD--EYIPSSSGPTGGGPKPDIVNIGAFEMASTR 663

233 SLAPDSEFWANDSKYAVNGGTSMAPIYAGNVAQLREHFVKNRGITPKPSLKAALLAG 292

664 TI-DGRGYGAQD-----VFGTSEATPYTSGTLALVFOAYKEVYNTTDPVTAKIILKSS 718

293 AADVGLGYPNGNQMGWRVTLDKSLNAVYNE 323

719 AKDI--WYPAFSGSGSRVDAKADIVFISE 747

RESULT 12

JC4908

alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.

N:Alternate names: subtilase

C:Species: Alteromonas sp.

C>Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 08-Oct-1999

A:Accession: JC4908

R:Unijob, H.; Miyamoto, K.; Tanaka, K.; Kaidzu, Y.; Imada, C.; Okami, Y.; Inamori, Y.

Biocell. Biotechnol. Biochem. 60, 1284-1288, 1996

A:Title: Cloning and sequence analysis of a protease-encoding gene from the marine bacterium

A:Reference number: JC4908; MUID:97141200; PMID:8987544

A:Accession: JC4908

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-715 <TSU>

A:Cross-references: DDBJ:D38600; NID:G1536787; PIDN:BAA18912.1; PID:d1019647; PID:g21602

A:Experimental source: strain O-7

C:Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-sensit

C:Genetics:

A:Gene: aprI

C:Superfamily: subtilisin homology

C:Keywords: hydrolase

F:1-40/Domain: signal sequence #status predicted <SIG>

F:41-150/Domain: amino-terminal propeptide #status predicted <ATP>

F:151-496/Product: alkaline serine protease I #status predicted <CAT>

F:182-452/Domain: subtilisin homology <SBT>

F:497-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F:239-394,335-372,478-461/Disulfide bonds: #status predicted

Query Match 12.0%; Score 270; DB 2; Length 715;

Best Local Similarity 26.3%; Pred. No. 9.1e-10;

Matches 126; Conservative 49; Mismatches 168; Indels 136; Gaps 23;

21 GGGQIVAVADTG---LDGRNDSMHEAFRGKITLXALGRNNMND-----TNG----- 67

182 GGGQVAVADTGTRPHLDLNLPGDMISNTFVANDGARDNDARDPDGAVTRECGT 241

68 -----HGTNVAG---SVLNGATNKGMAPQANLVQSIIMDSGGLGGLP 108

242 DSSGQPVPRADQDSWHTGTVAGTVAVTNNGGVAGVAYDAKVVPRVL---GKGGGLT 298

109 SNIQTLSQAFSAGARHTHTSKAAV-----NGAVTTDSRVNDVYKRNMTILFAAG 161

DB 299 SDIADGIIMASGSDRVPANAVAVINMSLGGGAGCSATQAINQANNGTVIYAG 358

162 NEGNNGGITSAPGAKNAITVGT--ENLPSFGSYADNINVAQSSSRGPTKGRIRKPD 219

359 NDNDNSANYN--FGNCQNVAVASVGDGSRAYSNGANI-----D 396

220 VMAPGTIILARSGLAPDSFWANHDS-----KYAVGTSMAPIYAGNVAQLR-- 269

399 VAPAG-----GAQSFADDPESGILSTHNSGSGAPENDSYHSQGTSMAPFAVAGVAAIIXQA 454

270 -----SHFVKN--RGITPKPSLKAALLAGADV--GLG-----YPNGNQMGWRVTL 312

455 KPSATPBEVERILKNTKRSFAGSCNCGTGVDAALVAVNEDLDVYTPPGN-----TL 508

313 DKSINAVYNESSALSTQKATYFTATAKPKLISLWSDAPASTASVTL---VNDD 369

509 ED--GVAKTGLSAGASGNQ--FFTFVPAK-----TNVFTMSGGTGDAD 550

370 LVITAPNGTXYGNDPAPEDNMWDR---NNVEVFINSPGSGTYTIEVQAYNVPV 424

551 LVYK-----LG--SQPTSSSTDCRPEGNAEVCSPAPQAGTYHWIINGYKAYSG 599

RESULT 13

A72647

probable surface layer-associated STABLE proteinase APE0607 - Aeropyrum pernix (strain

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: A72647

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Taka

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy

A:Reference number: A72450; MUID:99310339; PMID:10382566

A:Accession: A72647

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-1331 <XAM>

A:Cross-references: DDBJ:AP000060; NID:G5104188; PIDN:BAA79577.1; PID:d1043363; PID:g51

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0607

Query Match 11.6%; Score 260; DB 2; Length 1331;

Best Local Similarity 24.0%; Pred. No. 9.1e-09;

Matches 113; Conservative 67; Mismatches 153; Indels 138; Gaps 19;

8 VKADVAQSSVGLYGGQIVAVADTGDLTGRNDSMHE-ARFKITLVA----- 55

346 VYADLS-TAYTLF---LKALSDIGMTSGEPDPSLDDLSPADETPASYSGEVLARDFTD 400

56 -----LGRNNAN-----DTNGH 68

401 GVNDFSAALAGWTYDWVGLLTGESVNLGRLGFDYAGLVLPGLDPQGRWVSLIYDTLAI 460

69 GTHVAGSLTENG-----ATNKMAQANLVPSIIMDSGGLGGLPNSIQTLE-SQ 117

461 GTSVATVIASRGVNEVNLGYETSLRGVAPGAKI-----AAGG-----SFLINFAVQ 508

118 AFSAG-----ARHTNSWG---AAVNGAYT--TDSRVNDVY-VKRNDM 154

509 LFLSGFQSDPLMWVYTGHEQVDVINNSGNSYIALRGLTGADVDYATIEDIYVASGT 568

155 TILPAAGNEGNGTITAPGTAKAIVGATE--NLRPSGSAADNINVAQSSSRGPTK 212

569 VIVAMNGGQGYTATTPAGSLITSVGASTLFEDYVFPFGYVPSPGDVTISMDRSPQ 628

213 DGRITPDVMAPGTILSARSLAPDSFWANHDSKYAVMGTSMAPIYAGNVAQLREHF 272

629 IGVAKPDVNNIGSPAMNG---VPVLTLGNGSLAPDIPGTSBATPMSGSVALVISAY 684

273 VKRNGITPKPSLLAALIAAGADVGLGYPNGNQMGWRVTLDKSLNV-----A 319

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 7.51154 Seconds
(without alignments)
3008.498 Million cell updates/sec

Title: US-09-985-689a-2

Sequence: 1 NDVAGIVADVAQSSYGLY.....EVQATNVPGPNESLATVN 434

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	526.5	23.5	1743	1	TAGC_DICDI
2	499.5	21.8	1905	1	TAGC_DICDI
3	304.0	13.6	1398	1	PLS_PIRFU
4	302.5	13.5	806	1	SUBV_BACSU
5	278	12.4	420	1	SUBV_BACSU
6	277	12.3	580	1	EXPR_XANCP
7	259.5	11.6	513	1	ACLI_THBAO
8	250.5	11.2	534	1	PROX_VIBAL
9	250.5	11.2	402	1	ALP_CEPAC
10	249.5	11.1	894	1	WPR_BACSU
11	246	11.0	269	1	SUBS_BACLE
12	246	11.0	380	1	ELVA_BACAO
13	246	11.0	380	1	ELVA_BACCS
14	245.5	10.9	401	1	THES_BACSP
15	245	10.9	269	1	PRTW_BACSP
16	240	10.7	378	1	ELVA_BACSP
17	235	10.5	269	1	SUBS_BACLE
18	234.5	10.4	379	1	SUBT_BACLI
19	234.5	10.4	382	1	SUBT_BACAM
20	232.5	10.4	1052	1	MSLP_HUMAN
21	232.5	10.4	1181	1	SCAL_STRPY
22	232	10.3	331	1	ISP_BACCS
23	231.5	10.3	1052	1	MSLP_CRIGR
24	231.5	10.3	1052	1	MSLP_MOUSE
25	231.5	10.3	1052	1	MSLP_MOUSE
26	229.5	10.2	381	1	SUBN_BACNA
27	229	10.2	1167	1	SCAL_STRPY
28	227.5	10.1	381	1	SUBT_BACSU
29	225.5	10.0	275	1	SUBT_BACPU
30	225.5	10.0	381	1	SUBT_BACSU
31	225.5	10.0	381	1	SUBT_BACST
32	224	10.0	404	1	SMPI_MAGPO
33	219.5	9.8	274	1	SUBD_BACLI

34	219.5	9.8	645	1	SUBB_BACSU	P16396 bacillus su
35	218	9.7	279	1	THET_THEVU	P04072 thermoactin
36	217.5	9.7	361	1	ELVA_BACHD	P41363 bacillus ha
37	217	9.7	293	1	PRTT_TRIAL	P20015 tritrichia
38	217	9.7	422	1	TKSU_PYRKO	P58502 pyrococcus
39	214	9.5	326	1	ISP_FABPO	P29139 paenibacill
40	213	9.5	409	1	ALP_TRIHA	O03420 trichoderma
41	211	9.4	319	1	ISPI_BACSU	P11018 bacillus su
42	211	9.4	467	1	ISP6_SCHPO	P40903 schizosacch
43	210.5	9.4	388	1	CUDP_MERAN	P29138 metathizium
44	209.5	9.3	530	1	HLV_HALI7	P29143 halophilic
45	208.5	9.3	1902	1	P2P_LACPA	O02470 lactobacill

ALIGNMENTS

```

RESULT 1
ID TAGC_DICDI STANDARD; PRT; 1743 AA.
AC Q23868;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Prestalk-specific protein tagc precursor (EC 3.4.21.-).
GN TAGC
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=97140317; PubMed=8986798;
RA Shaulsky G., Escalante R., Loomis W.F.;
RT "Developmental signal transduction pathways uncovered by genetic
suppression."
RL Proc. Natl. Acad. Sci. U.S.A. 93:15260-15265(1996).
CC -!- FUNCTION: Intercellular communication via tagc may mediate
integration of cellular differentiation with morphogenesis (by
similarity).
CC -!- SIMILARITY: In the N-terminal section; belongs to peptidase family
S8.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING
TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -!- SIMILARITY: STRONG. TO TAGB.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation.
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U60086; AAB0331.1; -.
DR PIR; T18279; T18279.
DR DictyBase; DDB0001795; tagc.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PRF00664; ABC_membrane; 1.
DR Pfam; PRF00065; ABC_tran; 1.
DR Pfam; PRF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR SMART; SMO0382; AAA; 1.
DR PROSITE; PS50929; ABC_TMLF; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS00135; SUBTILASE_ASP; FALSE_NEG.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; FALSE_NEG.
DR HydroLase; Serine protease; ATP-binding; Transport; Transmembrane;
KW P00781 bacillus 11

```

KM	Signal		1	27	POTENTIAL.
FT	SIGNAL				PRESTALT-SPECIFIC PROTEIN TAGC.
FT	CHAIN	28	1743		
FT		316	6442		PROTASE.
FT	DOMAIN	1450	1687		ABC TRANSPORTER.
FT	DOMAIN	1450	1687		POTENTIAL.
FT	TRANSMEM	962	982		POTENTIAL.
FT	TRANSMEM	1027	1047		POTENTIAL.
FT	TRANSMEM	1157	1092		POTENTIAL.
FT	TRANSMEM	1260	1280		POTENTIAL.
FT	TRANSMEM	1288	1308		POTENTIAL.
FT	ACT SITE	325	325		CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT SITE	372	372		CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	637	637		CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	NP_BIND	1485	1492		AMP (POTENTIAL).
FT	DOMAIN	42	46		POLY-ASN.
FT	DOMAIN	94	103		POLY-ASN.
FT	DOMAIN	643	646		POLY-ALA.
FT	DOMAIN	733	741		POLY-ASN.
FT	DOMAIN	786	792		POLY-SER.
FT	DOMAIN	1337	1340		POLY-GLU.
FT	DOMAIN	1353	1352		POLY-GLY.
FT	DOMAIN	1353	1357		POLY-ASN.
FT	DOMAIN	1358	1364		POLY-ASP.
FT	DOMAIN	1361	1386		POLY-ASN.
FT	DOMAIN	1707	1729		POLY-ASN.
FT	CARBOHYD	390	390		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	536	536		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	547	547		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	614	614		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	669	669		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	735	735		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	741	741		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	776	776		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	832	832		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	887	887		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	1251	1251		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	1385	1385		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	1386	1386		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	1454	1454		N-LINKED (GLCNAC . .) (POTENTIAL)
FT	CARBOHYD	1704	1704		N-LINKED (GLCNAC . .) (POTENTIAL)
SEQ	SEQUENCE	1743 AA;	194145 MM;	12DB363E2F729639 CRC64;	
<hr/>					
Query Match		23.5%	Score 526.5;	DB 1;	Length 1743;
Best Local Similarity		27.8%	Pred. No. 9.6e-27;		
Matches 164;		Conservative 82;	Mismatches 152;	Indels 19;	Gaps 22;
<hr/>					
QY	19	LYGQCQIVAAVADTGIDTGR--NDS-----SMHEAFRGKITALYALGRTNNANDTNGH	68		
Db	314	LKGKQIILSIDTGTGDGSHCFPSDESKYPILPLNSVLNHR-KVTVITSTSDSDSKVDGH	372		
QY	69	GTHVAGSVLIG-----NGATNKGAPOANLVFGSIMSSGGGLGSLPNLOTLPFGAPSA	121		
Db	373	GHTIGSGAAGTFEDSDSVNISSSGLATDAKIFFLPLAGSSSLTF-PSDLQLGYPLPYDA	431		
QY	122	GARITNSWGA----AVNGAVYTDSRVNDVDYRKX-DMTILPAAGNEPGGFTIS--APG	174		
Db	432	GARYHCDSWGSVVEGYTSYESSDPASIDDFETFPHPDILLRAAGN---NEOVSILTLQS	486		
QY	175	TAKNAITVGATENTLR-----PSFGSYADN-----	199		
Db	469	TAKVITVTGAHQTIHENVLITDSPENVINVQSSVDINOELLCPDSTRXYCVTTAQCCLBESA	546		
QY	200	-----NHVAQSSSNQPTKGIKFDVMAFGTYILISA	230		
Db	549	TTGLASCCEPTLIRKSVIDAANTOPLLYNENNICSSSGKGFTHDGDMKPALVAFGEYITSA	608		
QY	231	RSSLA-----PDSSFMANHDISKVALYMAGTSVATPIVAGNYAQLEEH-----EVK	274		
Db	609	RSMGANNTDQCCDGL-PUTNALLA-ISGTSVATSPFAAATTILAQYLVDGVYPTPGSIVE	666		
QY	275	NRGITPKRSILKAALLTAG-----ADVGLGTPNGN-----QGMRKVTLDD	313		

```

Db      667  SNKLQPTGSLIKALMINNAQLNQTFLILTISSSTITFSNVQFENFAGASLVQGGALIRMS  726
QY      314  KSLVAVAYVNES-----ALSTSQATYTFY-----  338
Db      727  NWHLVNVNNNNNNNNKNTSDGITKEDG:GGDLRLVKPNQMKESLSTGQNTSYCFYKPS  786
QY      339  --ATACKPELK--ISLVSDAPASTASVTLVNDLDTLVITAPNGTRYVGNDFGAPEDNN  392
Db      787  SSSNSGNNIRFRVAVTLVMDDPPEYAAKFKPLVNNLLDTMI-----YYRNGSTIFYSN  840
QY      393  WDRG-----NNVENVFINSQSGTYTTEVQAVYVPCQFMS  429
Db      841  CGGSGSFGLAFTODTLNNVEGIVNTEPMTYRPMVAGTVWPGQNFMS  889

RESULT 2
TAGB_DICDI
ID  TAGB_DICDI  STANDARD;  PRT;  1905 AA.
AC  P54683;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  15-MAR-2004 (Rel. 43, Last annotation update)
DE  Prestalk-specific protein tagB precursor (EC 3.4.21.-).
GN  TAGB.
OS  Dictyostelium discoideum (Slime mold).
OC  Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
CX  NCBI_TaxID=44689;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=AX4;
RX  MEDLINE=95262903; PubMed=7744252;
RA  Shauleky G., Kuspa A., Loomis W.F.;
RT  "A multidrug resistance transporter/serine protease gene is required
RL  for prestalk specialization in Dictyostelium.";
RL  Genes dev. 9:1111-1121(1995)
CC  -1- FUNCTION: Intercellular communication via tagB may mediate
CC  integration of cellular differentiation with morphogenesis.
CC  -1- SIMILARITY: In the N-terminal section; belongs to peptidase family
CC  S8.
CC  -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING
CC  TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.
CC  -1- SIMILARITY: STRONG, TO TAGC.
CC  CC
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.jsb-sib.ch/announce/).
CC  or send an email to license@sb-ib.ch.
CC  CC
CC  EMBL; U20432; AAA6212.1; -.
DR  PIR; T18267; T18267.
DR  MEROPE; S08.UFM;-;
DR  DictyBase; DDB0001964; tagB.
DR  InterPro; IPR003593; AAA_ATPase.
DR  InterPro; IPR001140; ABC_TM_transp.
DR  InterPro; IPR003439; ABC_transporter.
DR  InterPro; IPR00209; Peptidase_S8.
DR  Pfam; PF00664; ABC_membrane_1.
DR  Pfam; PF00005; ABC_tran_1.
DR  Pfam; PF00082; Peptidase_S8; 1.
DR  PRINTS; PR00723; SUBTILISTIN.
DR  PRINTS; PD000006; ABC_transporter; 1.
DR  SMART; SMO0382; AAA; 1.
DR  PROSITE; PS50929; ABC_TM1F; 1.
DR  PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR  PROSITE; PS50693; ABC_TRANSPORTER_2; 1.
DR  PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
DR  PROSITE; PS00137; SUBTILASE_HIS; 1.
DR  PROSITE; PS00138; SUBTILASE_SER; 1.
DR  HydroLase; Serine protease; ATP-binding; Transport; Transmembrane;
KW  Signal.

```

```

FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 1905 PRESTALK-SPECIFIC PROTEIN TAG.
FT DOMAIN 376 700 PRETASE.
FT DOMAIN 1518 1756 ABC TRANSPORTER.
FT TRANSMEM 1011 1031 POTENTIAL.
FT TRANSMEM 1076 1096 POTENTIAL.
FT TRANSMEM 1121 1141 POTENTIAL.
FT TRANSMEM 1210 1230 POTENTIAL.
FT TRANSMEM 1309 1329 POTENTIAL.
FT TRANSMEM 1332 1352 POTENTIAL.
FT ACT SITE 387 387 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 432 432 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT SITE 695 695 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT NP_BIND 1553 1560 ATP (POTENTIAL).
FT DOMAIN 63 67 POLY-GLN.
FT DOMAIN 95 104 POLY-ASN.
FT DOMAIN 107 134 POLY-ASN.
FT DOMAIN 311 321 POLY-SER.
FT DOMAIN 833 837 POLY-SER.
FT DOMAIN 838 844 POLY-GLY.
FT DOMAIN 871 876 POLY-LEU.
FT DOMAIN 1012 1015 POLY-ILE.
FT DOMAIN 1386 1389 POLY-GLU.
FT DOMAIN 1398 1404 POLY-GLY.
FT DOMAIN 1445 1450 POLY-ASN.
FT DOMAIN 1765 1779 POLY-ASN.
FT DOMAIN 1782 1785 POLY-SER.
FT DOMAIN 1807 1812 POLY-PRO.
FT DOMAIN 1813 1860 POLY-PRO.
FT DOMAIN 1872 1878 POLY-PRO.
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 823 823 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1172 1172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1522 1522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1658 1658 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1905 AA; 212518 MW; B8E223FA8B9AE13C CRC64;

Query Match 21.8%; Score 490; DB 1; Length 1905;
Best Local Similarity 28.0%; Pred. No. 2,7e-24;
Matches 162; Conservative 74; Mismatches 159; Indels 184; Gaps 23;

QY 19 LVGGQGIYAVADTGLDPTGR---NDS-----SMHEAFGKLTALYALGRTNNDNTNGH 68
DB 376 LRGKQGLISTADTGLDGSHPFSDSKYPIPFNVNENHRYVT---YIYHNDYVNGH 432
QY 69 GTHVAGSVLANG---ATNKGAPOANLVFQSIMDSGGIGL-PSNLQTLFSQAFS 120
DB 433 GTHVCGSAAGTPEDSSMAISSFSGLATDKIAFYDL--SSGSGSEPTFPEDYSQMYKFLVD 490
QY 121 AGARHITSWGA---AVNGAYTTDSRNVDDYRK-NDMTILFFAAGNEGNGTISAPGT 175
DB 491 AGARHVHDSWGSVSLGGYGYGVDGAGIDAFIYEPERSILRAAGN-NELTASLLAQAT 549
QY 176 AKRAITVATENTRPSFGS-----YADNI----- 199
DB 550 AKRAITVGAEOIYAVYVSDALEYVDFSDANFORPCLFDKRYCNYTTAKCSESVNNGK 609
QY 200 -----NHVAQFSSRGPTDGRIRKPYVMAFGYIISARS 233
DB 610 LQLOCPASIKONASDSTTPOFYVENNNSSFSKPTTHDGRUKPIYAPGEYIISARSN 669
QY 234 -----LAPDSSWANHDSRYAMGTSWATPIVAGNVACLEBHF-----YKONG 277
DB 670 GENSTQCGDGSGL--PIANGIMISGTSWATPIATAATILIQYLVADGYFPTGSEVEERK 727
QY 278 ITRKPSLTKAAILAGADVGLG-----PENGQSGRATLDKSLNV 318
DB 728 LILPTGISLTKALMTNNALNGTYFWMASSTNPENALFEQINGANLILQMGALNAN--NW 784
QY 319 AYNNESS-----ALSTSQAKT-----YTFI-----ATAG 342

```

```

DB 785 LYVASSNPTPSRWIGIGLGNQKATEWKEDELSGLNKSICYFTYKPSSSSSGSGGGG 844
QY 343 KP-LKILVMSDPAASRTASVTIVNDL-----VITAPN--GTRVGNDFAPFD 350
DB 845 TPRVATLVMTDPSSYGAFNLVNNDDLILNSDDSIITIGSGSLQPAKVAP-- 902
QY 391 NMWDGRNNVENVPINSFQSGTYTIEVQAVNVYPGPONFS 429
DB 903 ---DTLNVGIIINPTKANNYKFTIAGTVNPIGPQKFS 938

RESULT 3
PLS_PYRPU STANDARD; PRT; 1398 AA.
AC P72186;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pyrolysin precursor (EC 3.4.21.-).
GN PLS OR P70287.
OS Pyrococcus furiosus.
OC Pyrococcus.
CC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
CX NCBI_TaxID=2261;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 150-184, AND CHARACTERIZATION.
RC STRAIN=Vci / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEDLINE=96355370; PubMed=8702780;
RA Voorhorst W.G.B., Eggen R.I.L., Geerling A.C.M., Platteeuw C.,
RA Sieszen R.J., de Vos W.M.;
RT "Isolation and characterization of the hyperthermostable serine
RT protease, pyrolysin, and its gene from the hyperthermophilic archaeon
RT Pyrococcus furiosus.";
RL J. Biol. Chem. 271:20426-20431(1996).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Vci / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weis R.B., Dunn D.M., Robb P.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome."
RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION, AND 3D-STRUCTURE MODELING.
RX MEDLINE=21079021; PubMed=11210516;
RA de Vos W.M., Voorhorst W.G.B., Dijkgraaf M., Kluskens L.D.,
RA Van der Oost J., Sieszen R.J.;
RT "Purification, characterization, and molecular modeling of pyrolysin
RT and other extracellular thermostable serine proteases from
RT hyperthermophilic microorganisms.";
RL Mech. Enzymol. 330:383-393(2001).
CC - FUNCTION: Has endopeptidase activity toward caseins, casein
CC fragments including alpha-S1-casein and synthetic peptides.
CC - SUBCELLULAR LOCATION: Cell envelope associated.
CC - PTM: LHM pyrolysin seems to be produced by autoprolytic
CC activation of HMW pyrolysin.
CC - PTM: Glycosylated.
CC - MISCELLANEOUS: Thermostable; high activity at 95 degrees Celsius.
CC - SIMILARITY: Belongs to peptidase family S8.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use. By non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DB EMBL: U55835; AAB09761.1;
DB EMBL: AE010153; AAL60411.1;
DB PIR: T28159; T28159.
DB HSSP: Q45670; ID51.
DB MEROPS: S08.100; -.
DB InterPro: IPR000209; Peptidase_S8.

```

InterPro: IPR007280; PPC.
 DR Pfam: PF00082; peptidase_s8, 1.
 DR Pfam: PF04151; PPC, 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE ASP, 1.
 DR PROSITE: PS00137; SUBTILASE HIS, 1.
 DR PROSITE: PS00138; SUBTILASE_SER, 1.
 KW Hydroxylase, serine protease; glycoprotein; zymogen; signal;
 complete proteome.
 FT SIGNAL 1
 FT PROPEP 26
 FT CHAIN 27
 FT ACT_SITE 150
 FT ACT_SITE 179
 FT ACT_SITE 365
 FT ACT_SITE 590
 FT CARBOHYD 152
 FT CARBOHYD 152
 FT CARBOHYD 222
 FT CARBOHYD 228
 FT CARBOHYD 240
 FT CARBOHYD 240
 FT CARBOHYD 257
 FT CARBOHYD 262
 FT CARBOHYD 298
 FT CARBOHYD 327
 FT CARBOHYD 406
 FT CARBOHYD 651
 FT CARBOHYD 651
 FT CARBOHYD 663
 FT CARBOHYD 739
 FT CARBOHYD 792
 FT CARBOHYD 893
 FT CARBOHYD 908
 FT CARBOHYD 917
 FT CARBOHYD 929
 FT CARBOHYD 1048
 FT CARBOHYD 1056
 FT CARBOHYD 1084
 FT CARBOHYD 1117
 FT CARBOHYD 1133
 FT CARBOHYD 1140
 FT CARBOHYD 1148
 FT CARBOHYD 1208
 FT CARBOHYD 1233
 FT CARBOHYD 1237
 FT CARBOHYD 1332
 FT CONFLICT 607
 FT CONFLICT 881
 FT CONFLICT 881
 SQ SEQUENCE 1398 AA; 154474 MW; 355D873A27D56582 CRC64;
 Query Match 13.68; Score 304.5; DB 1; Length 1396;
 Best Local Similarity 26.38; Pred. No. 2,7e-12;
 Matches 138; Conservative 57; Mismatches 177; Indels 153; Gaps 22;
 21 GGGGVAADADTGLDTRGDS-----SMHEAFNGKITALYALGRITNNAN----- 63
 301 GNGNDIYVDTDDYFTDEVPFGQVNTYDAVFSSYVGPLNYLAIEDPGEYAVFGM 360
 64 DTNGHGHVAGSYLNGATN-----KGNAPQAN 91
 361 DGHGHGHVAGSYLNGATN-----KGNAPQAN 91
 92 LVFQSLDSSGGIGLPSNLQTFQSAFSAARHNTNSGAAVNGAYT--TDSRN--VDD 147
 421 IMAIRVY-RSDGRGSMWDIEGM-TYATAGADIVANSIGG--NAYLIGTDPESVAVDE 476
 148 YVAKNDITLFAANGNEPNGGTSAPETAKNATLVGATENTLRPSFSYAD----- 197
 477 LTERKGVVFYIAANGSPGINIVGSPVATKALITVGAA--VPINAVIYSQALGYPDIYG 535
 198 -----NINHAQFSSRGPTDGRIRKPDVMAPGYIISARESLAPDSSFANHDSKYA 249
 536 FYFPAAYTNV-RIFAFSSRGPRIRIGEKPMVVAAGYIYSLSLPMWIGADF----- 585
 250 YMGTSWATPIVAGNVQLREHFVKNCGITPKXSLKALNLGAA-----DVGGLGYPNG 303

DB 586 -MSGTSWATPIVAGNVALLISG-AKAGGIYNDPIIKKVLSECATMLEGDPYGOXYTEL 643
 QY 304 NOGWRVTLDSLNVAVYNESSALSTSOKATYFTATNGKPLKISLWSDAPASTTA--- 360
 DB 644 DGHGHGHVAGSYLNGATN-----LKAINGTTLPIVDHMAKXSDFAEYL 685
 QY 361 SVTLVNDLDVITPKN-----GTRYVNDPFSAPD-----NNW-----DG-----RNVV--- 400
 DB 686 GVDVIRGLVARNISIPDIVEMHIKIVGDTERTREIYATEFPIKPVSGSVILANNTEFVL 745
 QY 401 -----NVFINSPOSGTYTIEVOAYNPVGPONFS 429
 DB 746 RVXVDVGLERGLVGRITLIDDP--TPVIEDLNTIVIPERFT 788
 RESULT 4
 SUBV_BACSU STANDARD; PRT; 806 AA.
 AC P29141;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Minor extracellular protease vpr precursor (EC 3.4.21.-).
 GN VPR OR IPA-45R OR BSU38090.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid=1423;
 [1]
 RP SEQUENCE FROM N.A. AND SEQUENCE OF 161-195.
 RX MEDLINE=92041574; PubMed=1938992;
 RA Sliema A., Ruto G.A. Jr., Theriault K.A., Dwyer M., Wilson S.W.,
 Pero J.;
 RT "Cloning and characterization of the gene for an additional
 RT extracellular serine protease of Bacillus subtilis";
 RT J. Bacteriol. 173:6869-6885(1991).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=95020537; PubMed=7934828;
 RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,
 Hullo M.F., Ionscu M., Lubochinsky B., Marcelino L., Moszer I.,
 Presecan E., Santanu M., Schneider B., Schweizer J., Vertes A.,
 Rapoport G., Danchin A.;
 RA "Bacillus subtilis genome project: cloning and sequencing of the 97
 RT kb region from 325 degrees to 333 degrees";
 RT Mol. Microbiol. 10:371-384(1993).
 [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=96044033; PubMed=9384377;
 RA Kunst F., Ogatawa N., Moszer I., Albertini A.M., Alloni G.,
 Azevedo V., Berteiro M.G., Bessieres P., Bolotin A., Borhett S.,
 Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 Brovilliet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
 Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 Denizot F., Devine K.M., Dusterhoft A., Enlich S.D., Emerson P.T.,
 Britan K.D., Errington J., Fabre C., Ferrazi E., Foulger D.,
 Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
 Ghm S.Y., Glaser P., Goffeau A., Golygity E.J., Grandi G.,
 Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,
 Kunita K., Lapilus A., Lardinois S., Lauber J., Lazarevic V.,
 Lee S.M., Levine A., Liu H., Maeda S., Manel C., Madigne C.,
 Medina N., Melisado R.P., Mizuno M., Moser D., Nakai S., Nodack M.,
 Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudga B., Park S.H.,
 Pato V., Pohl T.M., Portetle D., Porwolik S., Prescott A.M.,
 Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 Rieger M., Rivoita C., Rocha E., Roche B., Rose M., Sadie Y.,
 Sato T., Scanlan E., Schleich S., Schroeter R., Sciffone F.,
 Sekiguchi J., Sekowska A., Serrif S.J., Shin B.S., Solido B.,
 Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru T.,

RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosaio V., Uchiyama S., Vandenberg M., Vannier F., Vassaretti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zimstein E., Yoshikawa H., Darchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*,"
 RT Nature 390:249-256 (1997).
 CC -1- FUNCTION: NOT REQUIRED FOR GROWTH OR SPOULATION.
 CC -1- SUBCELLULAR LOCATION: C-TERMINAL.
 CC -1- PTM: PROBABLY UNDERGOES C-TERMINAL PROCESSING OR PROTEOLYSIS.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M76590; AAA2881.1; -
 DR EMBL; X73124; CA451601.1; -
 DR EMBL; Z39123; CA315835.1; -
 DR PIR; A41341; A41341.
 DR HSSP; P00783; 2S8T.
 DR MEROPS; S08.00A; -
 DR Subtilisin; BGI0591; VPR.
 DR InterPro; IPR001373; PA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR009020; Protease_inhib.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF02225; PA; 1.
 DR PRINTS; PRO0723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILISIN_ASP; 1.
 DR PROSITE; PS00137; SUBTILISIN_HIS; 1.
 DR PROSITE; PS00138; SUBTILISIN_SER; 1.
 KM Hydrolyase; Serine protease; Zymogen; Signal; Complete proteome.
 FT SIGNAL; 1 28
 FT PROPEP; 29 160
 FT CHAIN; 161 806
 FT ACT_SITE; 189 189
 FT ACT_SITE; 233 233
 FT ACT_SITE; 534 534
 SQ SEQUENCE 806 AA; 85608 MW; P984E3BF0869DDC CRC64;
 Query Match 13.5%; Score 302.5; DB 1; Length 806;
 Best Local Similarity 23.2%; Pred. No. 1,8e-12;
 Matches 127; Conservative 58; Mismatches 161; Indels 201; Gaps 17;
 QY 18 GLYGQGIIVAVADTGLDTR-----NDSSVHEAFRCKITLALVALGRTN 60
 DB 177 GYRGKGIKVALIDTGEVNHPRDKNFGYGYDVFNDYDKERTPTG-----D 225
 QY 61 NAADTNGHGHVAGSLTNGGATKKAAPQANLVFQSSGLDGLPSNLTLSQAPS 120
 DB 226 PRGEATDGHVAGVLAANG-TIKGVAPEATLLAVRLPQGG-SGTENVAVGERAVQ 282
 QY 121 AGARHTNSWGAANVAYTTDSRNVDYRKNDMTILFAAGNEGPGGATISAPTAQAI 180
 DB 283 DGDVNVNLSLGNLNPDMATSTAL-DWAMSEGVAVYNSNGSNGMVGSPGSRRAI 341
 QY 181 TVQATE---NLRFPSGGSY----- 195
 DB 342 SVGATQPLNEVAVTGGSYSAKVGYNKEDDVKALNNKEVELVEAGIGEAQDFGKDLT 401
 QY 196 -----ADNI----- 199
 DB 402 GKVAIVKRGSIATVDKADNKAAGATGMVYNNLSGEIAPNPGMSVPTIKSLDEGKL 461
 QY 200 -----NHVAFSSRGPTKD-GIKPDVAPGTYIILSARSSLAP 236
 DB 462 VSALKAGETKTPKLTIVSKALGEQVADFPSSRGVMDTWIKPDISAPGVNIIVSTIPTHDP 521

QY 237 DSSFWANHDSKAYMGSTSWATPIVAGNVAQAREHFNKRGITPKPSL--IKKALTAGA 294
 DB 522 D-----HPYIGSKQSTSNASHPHLKAIVAKO-----AKPKASVEGIRKALINNTAY 568
 QY 295 DV-----GLGYPNGNQNGVRYTLDRKSNVAVYNESSALSTQKATYFTATAGKPLKISLV 350
 DB 569 TLKSDSEVYPHNAQAGASARI---MVA--IKADSLVSPGSYSGFTLKENGETKNETF 623
 QY 351 WSDAPASTASVTLVNDLVIATPAGTRVGNDFGAPFNMDGNANNVFINPSQSG 410
 DB 624 TIEQSSIRKSYTL-----EYS--FVSGSISTSGISRVVIPAHTQ 662
 QY 411 TYTIEVQ 417
 DB 663 KATKVK 669
 RESULT 5
 SUBT_BACS9 STANDARD; PRT; 420 AA.
 AC P28842;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Subtilisin precursor (EC 3.4.21.62).
 GN SUBI.
 OS *Bacillus* sp. (strain TA39).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; *Bacillus*.
 OX NCBI_TaxID=29336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92256481; PubMed=1581352;
 RA Narix E., Davail S., Feller G., Gerday C.;
 RT "Nucleotide and derived amino acid sequence of the subtilisin from
 RL the antarctic psychrotrophic *Bacillus* TA39,"
 RL Biochem. Biophys. Acta 1131:111-113 (1992).
 CC -1- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
 CC it catalyzes the hydrolysis of proteins and peptide amides.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
 CC for peptide bonds, and a preference for a large uncharged residue
 CC in P1. Hydrolyzes peptide amides.
 CC -1- COFACTOR: Binds 1 calcium ion per subunit (potential).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: Still active at temperatures close to 0 degrees
 CC Celsius; it has a marked heat lability.
 CC -1- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
 CC sporulation, and many mutations which block sporulation at early
 CC stages affect expression levels of subtilisin. However, subtilisin
 CC is not necessary for normal sporulation.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X62369; CA444227.1; -
 DR PIR; S23407; S23407.
 DR HSSP; Q99405; 1WPT.
 DR MEROPS; S08.00A; -
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PRO0723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILISIN_ASP; 1.
 DR PROSITE; PS00137; SUBTILISIN_HIS; 1.
 DR PROSITE; PS00138; SUBTILISIN_SER; 1.
 KM Hydrolyase; Sporulation; Serine protease; Zymogen; Metal-binding;
 FW Calcium-binding; Signal.
 FT SIGNAL 1 26
 FT SIGNAL 26
 POTENTIAL.

```

FT PROPEP 27 111 POTENTIAL.
FT CHAIN 112 420 SUBSTITISIN.
FT ACT_SITE 145 145 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 182 182 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 360 360 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT METAL 115 115 CALCIUM (POTENTIAL).
FT METAL 154 154 CALCIUM (POTENTIAL).
SQ SEQUENCE 420 AA; 44086 MM; AE4F21BD32B266C CRC64;

Query Match 12.4%; Score 278; DB 1; Length 420;
Best Local Similarity 33.2%; Pred. No. 3, 1e-11;
Matches 88; Conservative 38; Mismatches 95; Indels 44; Gaps 13;

QY 21 GGGQIVAVADTGLDTRGNDSSMHEAFRGKITLAL--VALGRT--NNANDNGCTVAGS 75
Db 136 GGGININAVLDGTGVTN-----HDLRANVEOCKDFVGTYYTNNSCDTRQCGHVAAS 189
QY 76 VLNGATNK--GNAPQANLVFQSIM--DSSGGLGGLPSNLQTLFQAFSAGARIMTN-S 129
Db 190 ALADGGTGNVGVAVPADIMAYKVLGDGSGYADDIAPAIIRAGQATRLTKVYINMS 249
QY 130 MGAAVNAYTTDSNVVDYRKNDMTLLFAAGNEPRNGCTISAPGAKAHITGATENR 189
Db 250 LGSSGESSLITNAN--YSYNGVLLTAAAGNSGPQSGTGYGALVNAVAALLEN-K 305
QY 190 PSEGSYADNINHAQFSSRGPT-KDG-----RIKPDVMAFGTYILSARSLAPDSSFWAN 243
Db 306 VENGTY-----RVADFSRGSWTDGDYAIQKGDVEISAGALVST-----W-- 348
QY 244 HDSKYAYMGTSMTATPVAGNVAQL 268
Db 349 PDGGYATITSGTSMASPHAGLAAT 373

RESULT 6
EXPR_XANCP STANDARD; PRT; 580 AA.
ID EXPR_XANCP
AC P3314;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Extracellular protease precursor (EC 3.4.21.-).
GN XCC0851.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
CX NCBI_taxid=340;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90251253; Pubmed=2187155;
RA Liu Y.-N., Tang J.-L., Clarke B.R., Dow J.M., Daniels M.J.;
RT "A multipurpose broad host range cloning vector and its use to
RT characterise an extracellular protease gene of Xanthomonas campestris
RT pathovar campestris";
RL Mol. Gen. Genet. 220:433-440(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33913 / NCPB 528;
RX MEDLINE=2022145; Pubmed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach P.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Camarotte G., Canabarro F., Cardozo J., Chambeiro F., Clapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Mendes J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.C., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spindola L.A.F., Takita M.A., Tamura R.B., Teixeira B.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,

```

```

RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities";
RL Nature 417:459-463(2002).
CC -|- SUBCELLULAR LOCATION: Secreted.
CC -|- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; X51635; CA35962.1; -.
CC DR EMBL; AE012184; AA040166.1; -.
CC DR PIR; S11890; S11890.
CC DR HSSP; P00782; 2SPT.
CC DR MEROPS; S08.0PA; -.
CC DR InterPro; IPR000209; Peptidase_S8.
CC DR InterPro; IPR007280; PC.
CC DR InterPro; IPR009020; Protease_inhib.
CC DR Pfam; PF00082; Peptidase_S8; 1.
CC DR Pfam; PF04151; PPC; 1.
CC DR PRINTS; PR00723; SUBSTITISIN.
CC DR PROSITE; PS00136; SUBSTITISIN_ASP; 1.
CC DR PROSITE; PS00137; SUBSTITISIN_HIS; 1.
CC DR PROSITE; PS00138; SUBSTITISIN_SER; 1.
CC DR KMW; K00138; Serine protease; Zymogen; Signal; Complete proteome.
CC FT SIGNAL; 1 32
CC FT PROPEP; 33 7136
CC FT CHAIN; 7137 580
CC FT ACT_SITE; 177 177
CC FT ACT_SITE; 237 237
CC FT ACT_SITE; 409 409
CC FT ACT_SITE; 425 425
CC FT DISULFID; 315 352
CC FT DISULFID; 450 454
CC SQ SEQUENCE 580 AA; 57228 MM; 8C9A2ACAB4E7F47CB CRC64;

Query Match 12.3%; Score 277; DB 1; Length 580;
Best Local Similarity 25.6%; Pred. No. 5, 5e-11;
Matches 126; Conservative 53; Mismatches 150; Indels 144; Gaps 22;

QY 21 GGGQIVAVADTGL-----DTGRNDSMHEAFRGKITLALYALGRTNAND----- 64
Db 168 GSGTVAVAVDITGIRSHDLNANILAGYDPLSDATTARDNGRNSNADBDGWTAAABCGA 227
QY 65 -----TNGHGTHTAGSVLNGATNKGMAPQ-----NLVFSIMDS 100
Db 228 GIPASSSWHGTHTAGVAVAVNTTGTAGTAAGAKVPPVILGCKGSLSDADALVMA 287
QY 101 SGG-LGGLPSNLQ--TFQAFSAGARIHTNSMGAAVNGAYTTDSRVVDYRKNDMTIL 157
Db 288 SGGTVSGIPANANAEVYINNSLGGGSCSTTMQN-AINGAVSNGT-----TVV 334
QY 158 FAAGNEBNGGTIAPGTAKNAITVGTEN--LRPSGSAADNINHAQFSSRGPTDCGR 215
Db 335 VAAANDASNVSG-SLPANCANVIAVAATTSAGAKASVSNGTGI----- 377
QY 216 IKPVMVMAFGTYILSARSS--LAPDSSFWANHDSKYAYMGTSMTATPVAGNVAQLREHFV 273
Db 378 ---DVSAFGSSILSTLNSGTTTPS-----ASYASYNSTSMASPHVAVGVALVYQS--V 425
QY 274 KNRGITPK--PSLIK--AALAGADVGLG-----PENGQSGRVTLDK 314
Db 426 APTALTTPAAVETLLKNTRARLPAGSGCGAGIVNDAVTAALINGSGGSGGAGNTLIN 485
QY 315 SLNAYVNESAISTGKATYTFATGKPKLISLWSDAPASTGASVTL---VNDPDLV 371
Db 486 GTPVT-----GIGATGALNVTITVPAG-----SGTLVTTSSGSGGADLV 527

```


QY 372 ITAPNGRYVNDSPAFDNNMDGR-----NNVENVFINSPOSQTYTTEVOAYN 420
 DB 528 VRAQ-----SAPTDSATYCRPRPSGNATCTCTTAP-SQTYVRLKAYS 569

RESULT 7
 AQLI_THEAQ STANDARD; PRT; 513 AA.
 AC P08594;
 DT 01-AUG-1998 (Rel. 08, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Aqualysin I precursor (EC 3.4.21.-).
 GN PSTI.
 OS Thermus aquaticus.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
 OC Thermus.
 NCBI_TaxID=271;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 15-23.
 RC STRAIN=YT1;
 RX MEDLINE=9021674; PubMed=2182621;
 RA Terada I., Kwon S.-T., Miyata Y., Matsuzawa H., Ohta T.;
 RT "Nucleotide sequence of an extracellular protease, aqualysin
 I with NH2- and COOH-terminal pro-sequences and its processing in
 Escherichia coli.";
 RL J. Biol. Chem. 265:16576-6581(1990).
 RN [2]
 RP SEQUENCE OF 75-442 FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=YT1;
 RX MEDLINE=88225062; PubMed=3286255;
 RA Kwon S.-T., Terada I., Matsuzawa H., Ohta T.;
 RT "Nucleotide sequence of the gene for aqualysin I (a thermophilic
 alkaline serine protease) of Thermus aquaticus YT-1 and
 characteristics of the deduced primary structure of the enzyme.";
 RL Eur. J. Biochem. 173:491-497(1988).
 RN [3]
 RP SEQUENCE OF 128-170.
 RX MEDLINE=8815937; PubMed=3162211;
 RA Matsuzawa H., Tokugawa K., Hamachi M., Mizoguchi M., Taguchi H.,
 Terada I., Kwon S.-T., Ohta T.;
 RT "Purification and characterization of aqualysin I (a thermophilic
 alkaline serine protease) produced by Thermus aquaticus YT-1.";
 RL Eur. J. Biochem. 171:441-447(1988).
 CC -1- FUNCTION: Aqualysin I is a thermophilic alkaline serine protease.
 The optimal temperature for its caseinolytic activity is 80
 degrees Celsius.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- DEVELOPMENTAL STAGE: Secreted from the early stationary phase
 until the time the cells cease to grow.
 CC -1- PTM: THE N- AND C-TERMINAL PRO-SEQUENCES ARE REMOVED THROUGH THE
 PROTEOLYTIC ACTIVITY OF AQUALYSIN I ITSELF, IN THAT ORDER. THE
 C-TERMINAL PRO-SEQUENCE IS REQUIRED FOR TRANSLOCATION OF THE
 PROTEIN ACROSS THE OUTER MEMBRANE.
 CC -1- PTM: Two disulfide bonds are present.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL/ D90108; BAAL135.1; -
 DR EMBL/ X07734; CA30559.1; -
 DR PIR/ A35742; A35742.
 DR HSSP/ P06873; 2PRK.
 DR MEROPS/ S08.051; -
 DR InterPro/ IPR000209; Peptidase_S8.
 DR InterPro/ IPR009020; Protease_inhib.
 DR Pfam/ PF00082; Peptidase_S8; 1.

DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Hydroxylase; Serine protease; Zymogen; Signal.
 FT SIGNAL 1 14
 FT PROPEP 15 127
 FT CHAIN 128 408
 FT PROPEP 409 513
 FT PROPEP 409 513
 FT ACT SITE 166 166
 FT ACT SITE 197 197
 FT ACT SITE 349 349
 FT ACT SITE 349 349
 SQ SEQUENCE 513 AA; 53913 MW; DDFDFED4A50B785 CRC64;

Query Match 11.6%; Score 259.5; DB 1; Length 513;
 Best Local Similarity 26.4%; Pred. No. 6.6e-10;
 Matches 119; Conservative 45; Mismatches 141; Indels 145; Gaps 24;

QY 16 SYGLYGGQGLVAVADTGLDTRNDSSMHEAFRGKITLY-ALGRTNNANDTNGGTHVAG 74
 DB 152 TYTAGRGVAVYVIDTGIRT-----THREFGRARVGYDALG--GNGQDCNGGTHVAG 203
 QY 75 SYLGNGATNGMAQANLVFQSIMD--SSGGLGLPBNLOTLSQAFSAGARTHTN----- 128
 DB 204 TI--GGVTVGAVALVAVRVIDCNGSGSTSGVLAGVWV-----TRNHRPPAVA 252
 QY 129 --SWGAAVNGAYTTDSNVVDYVRKN--DMTILFAAGNEGPNGGTISAPGTAKNAITVG 183
 DB 253 NMSLGGGVSTA-----LNAVAVNSIAAGVYVAAGNNANAVNS-PARVAELITVG 304
 QY 184 AT--ENLRPSFGSYADINNVAVQSSRGPKDGRKIPDVAPRTYILISAASSLAPDSFW 241
 DB 305 ATTSDDPARSSNNGSCV-----DLFAPGASIPSA-----W 335
 QY 242 AMHDSKAYMGSTMATPIVAGVAVQUREHFVKNRGITP---XPSLIKALLIAGADVGL 298
 DB 336 YTSPTATQTLNGTSMATPHVAG--VAAL--YLEQNPSTATPASVASAILNGTTGRLSGIGS 392
 QY 299 GYPNGNGMGKRLIDKLANVAYNNESSALSTSCAKYTFPATAGKPKILSLVSDAPAST 358
 DB 393 GSPN-----RLLY-----SLSSGSGST-----APCTS 415
 QY 359 ---TASVTVLNDLDVITAPNGTRY-----VGNDSPAFDNNMDGRN--- 397
 DB 416 CSYITGSLSGGDVNF--QPNGTYYISPACTHRAMLRGAGTDFDL-YIMRDGSRMLT 471
 QY 398 -----NVENVFINSPOSQTYTTEVOAYN 420
 DB 472 VGSSTGPTSEBSLSYSTGAGYIMRIYAYS 501

RESULT 8
 PROA_VIBAL STANDARD; PRT; 534 AA.
 AC P16588;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alkaline serine exoprotease A precursor (EC 3.4.21.-).
 GN PROA.
 OS Vibrio alginolyticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 NCBI_TaxID=663;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89325126; PubMed=2546861;
 RA Deane S.M., Robb F.T., Robb S.M., Woods D.R.;
 RT "Nucleotide sequence of the Vibrio alginolyticus calcium-dependent,
 detergent-resistant alkaline serine exoprotease A.";
 RL Gene 76:281-288(1989).
 CC -1- SIMILARITY: Belongs to peptidase family S8.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M25499; AAA27550.1; -;
 DR PIR; J50173; J50173.
 DR HSPB; Q99405; 1MPT.
 DR MEROPS; S08.050; -;
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR007280; PRC
 DR InterPro; IPR009020; Protease_inhib.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PRC; 1.
 DR PRINTS; PRO0723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydroxylase; Serine protease; Zymogen; Signal.
 FT SIGNAL 1 21
 FT PROPEP 22 141
 FT CHAIN 142 534 ALKALINE SERINE EXOPROTEASE A.
 FT ACT_SITE 180 180 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 363 363 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 534 AA; 55930 MW; 84E96D9C649D4226 CRC64;

Query Match 11.2%; Score 252.5; DB 1; Length 534;
 Best Local Similarity 24.3%; Pred. No. 2e-09;
 Matches 116; Conservative 70; Mismatches 137; Indels 155; Gaps 25;

QY 7 IVKADVAGSS--YGL-----YGGQGVAVAVDTGLDTRNNSHEAF 46
 DB 137 IVSDANQTNAINKLDIDQNLPLNNYANFPGTGVTVYVDTGV-----NNHVEF 190
 QY 47 RGTITALY-ALGRTNANDTNGHAGSVLNGATNKMAPQANLVFQSI--DSGG 103
 DB 191 GGRSVSGVDFYDNDADSDCGHGTVAAGTI---GSLYVAKNVLVGRVLSGSGS 247
 QY 104 LGG-PSNLQTLFSGAFAGARIHTNSGAVNGAYTTDSUNVDYR---KNMTLLEFA 160
 DB 248 TSGVTAQVDAVA--SGPSVANSLGCGQVA-----LDSAVQSAVQSGVSFLLA 298
 QY 161 GNEGPNQGTISAPGTAKAATTVGAT--ENDRPSFGSYADININEVAQFSSKPTKDIRKP 218
 DB 299 GNSVADACNYS-PARVATGVTVGTTSTIDARSSFSMWGSCV----- 338
 QY 219 DYMAFGVYILSARSLAPDSSFMANHSKAYMGSTSMATPIYAGNVAQIREHFKNGRI 278
 DB 339 DVFPFGSQIISA-----W-YDGGYKTIISGTSMAFHVAG--VAAL--YLOENSSV 383
 QY 279 TPKPSLKAALIAADAVGLGYPNQGWGRVTLDSKINAYVNESSALSTOKATYTF 338
 DB 384 S--PSQVEALIVSA--GKVT-----DTGSGVVKLXLSLT 416
 QY 339 -----ATAGKPL-----KISLVMSAPASTTASVTLV--ND 367
 DB 417 DADCGQCGGDPPTPEDEGKLTISGVPSGLSGSGGVAAYVYVEGQRLTVQMYGSGSD 476
 QY 368 LDVITAPNGTRVYGNDFSA-PEDNNMDGR-----NNVENVFINSPOSGTYTIEVCAYN 420
 DB 477 ADVLIR-----FGAKPTLNAMDGRPKYGNNETCTVSAITSGRHHWIGQYS 523

RESULT 9
 ID_CEPAC STANDARD; PRT; 402 AA.
 AC P29118;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Alkaline proteinase precursor (EC 3.4.21.-) (ALP).
 GN ALP.
 OS Cephalosporium acremonium (Acremonium chrysogenum).
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Hypocremycetidae; Hypocreales; Hypocreaceae; mitosporic Hypocreaceae;
 CC Acremonium.
 OK NCBI_TaxID=5044;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=91299283; PubMed=1368696;
 RA Isogai T., Fukagawa M., Kojima M., Aoki H., Tanaka H.;
 RT "Cloning and nucleotide sequences of the complementary and genomic
 RL DNA for the alkaline protease from Acremonium chrysogenum.";
 RL Agric. Biol. Chem. 55:471-477(1991).
 CC -1- SIMILARITY: Belongs to peptidase family S8.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D00923; BAA00765.1; -;
 DR PIR; J00332; J00332.
 DR HSPB; P06873; 2PRK.
 DR MEROPS; S08.DPA; -;
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR009020; Protease_inhib.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PRO0723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydroxylase; Serine protease; Zymogen; Signal.
 FT SIGNAL 1 20
 FT PROPEP 21 120
 FT CHAIN 121 402 ALKALINE PROTEINASE.
 FT ACT_SITE 160 160 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 191 191 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 347 347 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 402 AA; 42099 MW; BD030CCD42D918E1 CRC64;

Query Match 11.2%; Score 250.5; DB 1; Length 402;
 Best Local Similarity 31.0%; Pred. No. 1.9e-09;
 Matches 93; Conservative 31; Mismatches 101; Indels 75; Gaps 14;

QY 21 GGGQGVAVAVDTGLTGRNDSMEHAFRGK-ITALYALGRTNANDTNGHGTVAAGSVLGN 79
 DB 151 GSGTYAVVVDVGT-----LESHNEFSGRAITGVNAGGSEN--ADTNGHGTVAAGTI--- 199
 QY 80 GATNKMAPQANLVFQSIIMSSGQ-----LGLPSNLQTLFSGAFSAAGARI-----HT 127
 DB 200 GGRVGVAKNTNLAVVFRSSSSSTSIILDFPMVAVMDIINRQKKAISLSLGGQYS 259
 QY 128 NSGAVNGAYTTTSKAVVDYRKNDMTILFAAGNEPNQGTISAPGTAKAATTVG--AT 185
 DB 260 SAFNNAVNTAY---SRGLSVY-----AANDNQANYS-PASANAATVGSIAS 306
 QY 186 ENLRPSFGSYADININHAQFSSRPTDGRIKPVMAPFGVYILSARSLAPDSSFMANHD 245
 DB 307 NMASSFENYGSV-----DIFAPGISISA-----WIGCN 337
 QY 246 SKYAVMGSTSMATPIYAGNVAQIREHFKNGRIKPKPSLKAALIAADAVY--GLGYEN 302
 DB 338 SATNTISGTSMAFHVGLVLIQ-----ALEGLTSGAARLNLNLAATTGTVSNFGSGSPN 393

RESULT 10
 WPIRA_BACSU STANDARD; PRT; 894 AA.
 ID WPIRA_BACSU

CC P54423; 006726;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cell wall-associated protease precursor (EC 3.4.21.-) [Contains: Cell
 wall-associated polypeptides CWBP23 and CWBP52].
 GN WPRA OR BSU10770.
 OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OK NCBI_TaxID=1423;
 RN 11
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54 AND 414-428.
 RC STRAIN=168;
 RC MEDLINE=97158234; PubMed=9004506;
 RA Margot P., Karamata D.;
 RT "The wprA gene of Bacillus subtilis 168, expressed during exponential
 growth, encodes a cell-wall-associated protease.";
 RL Microbiology 142:3437-3444(1996).
 RN 12
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RC MEDLINE=98015415; PubMed=9353931;
 RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;
 RT "Sequencing of regions downstream of addA (98 degrees) and citG (289
 degrees) in Bacillus subtilis.";
 RL Microbiology 143:3305-3308(1997).
 RN 13
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RC MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moser I., Albertini A.W., Alloni G.,
 RA Aaeved V., Bertero M.G., Bessieres P., Bolotin A., Borcher S.,
 RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Brusch C.V., Caldwell B., Capriano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Deniot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Ertter K.D., Errington J., Fabret C., Ferrati E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizi A., Galleron N.,
 RA Gilm S.Y., Glaeser P., Goffeau A., Goldschly E.J., Grandi G.,
 RA Giallippi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Rauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Maeda S., Manuel C., Medgyes V.,
 RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawa A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,
 RA Sekiuchi U., Sekowska A., Seror S.J., Seror P., Shin B.S., Solido B.,
 RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takemaru K.,
 RA Takenchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Toso V., Uchiyama S., Vandenbol M., Vannier F., Vassonelli A.,
 RA Vairi A., Wambitt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zuretein E., Yoshikawa H., Darchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.";
 RL Nature 390:249-256(1997).
 CC - FUNCTION: NOT YET KNOWN; COULD BE INVOLVED IN PROTEOLYTIC
 CC DEGRADATION, BY CLEAVAGE OF ITS PEPTIDE BRIDGES.
 CC - SUBCELLULAR LOCATION: Cell-wall bound.
 CC - PTM: PROCESSED INTO CWBP23 AND CWBP52.
 CC - SIMILARITY: Belongs to peptidase family S8.

CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U58981; AAC25926.1; -
 DR EMBL; Y09476; CA17064.1; -
 DR EMBL; Z99109; CAB12917.1; -
 DR PIR; F69730; F69730.
 DR HSP; Q45670; IDBI.
 DR MER05; S08.004; -
 DR InterPro; Bg1846; WPRA.
 DR InterPro; IPR00209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR HydroLase; Serine protease; Cell wall; Zymogen; Signal;
 KM Complete proteome.
 FT SIGNAL 1 31
 FT CHAIN 32 894
 FT PROPEP 32 413
 FT CHAIN 414 894
 FT ACT_SITE 466 466
 FT ACT_SITE 497 497
 FT ACT_SITE 650 650
 FT CONFLICT 9 9
 FT CONFLICT 14 14
 FT SEQUENCE 894 AA; 96487 MW; 0F67C353E5F8DBC CRC64;
 Query Match 11.18; Score 249.5; DB 1; Length 894;
 Best Local Similarity 24.28; Pred. No. 6.1e-09;
 Matches 106; Conservative 63; Mismatches 140; Indels 129; Gaps 19;
 CC -----
 QY 25 IVAAVDGLDTRGRNDSMEHAFGKITALYA---LGRTNNAQNTNGHGHVAGSVG--- 78
 DB 457 LIAVVDGVSTLAD-----LKKVRTDIGHFVGRRNNAMDDQGHGHTVAGIIAASD 510
 QY 79 NGATNKMADQANLVFQSIWSSGGLGSPNQLTFSQAFSAARLHTNSWGAANGAY 138
 DB 511 NGYSMTGLNKKAKILPKYVLDSDAG--SGDTEQIALGIKYAADGAKVIMSLG---GGV 564
 QY 139 TTDSRNVDDVVRKNDMLFLAAGNEGNGTISAPGAKAIVGATENLRPSFGSYADN 198
 DB 565 SVLEPLKTKAADKRVILAAASGNDEN--ALSYPSKTVSGVAT-----NR 611
 QY 199 INHVAQSSRGPTDGRKEDVVAFGTYILSARSSLPASSFPVANDSKYAVYGTSMAT 258
 DB 612 MDMTADPSNYGKL-----DISAPGSDI---PSLVNNGN-----VTYVSGTSMAT 653
 QY 259 PIVAGNVAGLRHENVKRGIL--TPKSLKALAIAGADVGLGVPGNQ----- 305
 DB 654 PYAAAAAGLL--FAQPKTKTEVEDMLKKT---ADTISFSSVGGEBELYDDYDGP 706
 QY 306 -----GNGRYTLDKSLNVA---YVNESSALSTSQKAT-----Y 335
 DB 707 EIKPTGVMDHSGVGRNWKAVASADLQKVKKLSGTQYAVGSAKEGTLLIEMNGKK 766
 QY 336 TETITAK--PLKISLYMSAPASTASVTLVNDLDLVITAPNGTRYGVNDSAPDNNW 353
 DB 767 LGSAAKAKDNNAFKNI-----ATQKQDQVLYLKATRG----- 798
 QY 394 DGRNNEVNFINSQSGT 411
 DB 799 DAKTSYVVVVKKGPST 816
 CC -----
 RESULT 11
 SUBS_BACLE STANDARD; PRT; 269 AA.
 AC P29600;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Subtilisin Savinase (EC 3.4.21.62) (Alkaline protease).
 OS Bacillus lentus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1467;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
 RX MEDLINE=92148829; PubMed=1738156;
 RA Betzel C., Klupsch S., Papendorf G., Hastrup S., Branner S.,
 Wilson K.S.;
 RT "Crystal structure of the alkaline proteinase Savinase from Bacillus
 lentus at 1.4-A resolution."
 RJ J. Mol. Biol. 223:427-445(1992).
 RL [2]
 RN STRUCTURE BY NMR.
 RX MEDLINE=96184541; PubMed=8654411;
 RA Remerowski M.L., Pepermans H.A.M., Hilbers C.W., van de Ven F.J.M.;
 RT "Backbone dynamics of the 269-residue protease Savinase determined
 from 15N-NMR relaxation measurements."
 RJ Eur. J. Biochem. 235:629-640(1996).
 RL [3]
 RN X-RAY CRYSTALLOGRAPHY (0.78 ANGSTROMS).
 RX MEDLINE=98426039; PubMed=9753430;
 RA Kuhn P., Knapp M., Soltis S.M., Ganshaw G., Thoenes M., Bott R.;
 RT "The 0.78-A structure of a serine protease: Bacillus lentus
 subtilisin."
 RJ Biochemistry 37:13446-13452(1998).
 RL [1]
 CC -1- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
 it catalyzes the hydrolysis of proteins and peptide amides.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
 for peptide bonds, and a preference for a large uncharged residue
 in P1. Hydrolyzes peptide amides.
 CC -1- COFACTOR: Binds 2 calcium ions per subunit.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- BIOTECHNOLOGY: Used as a detergent protease. Sold under the name
 savinase by Novozymes.
 CC -1- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
 sporulation, and many mutations which block sporulation at early
 stages affect expression levels of subtilisin. However, subtilisin
 is not necessary for normal sporulation.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 DR PDB: 1C9J; 06-OCT-99.
 DR PDB: 1C9N; 10-JAN-01.
 DR PDB: 1C9M; 10-JAN-01.
 DR PDB: 1GCI; 11-NOV-98.
 DR PDB: 1IAV; 11-JUL-01.
 DR PDB: 1JEA; 26-NOV-97.
 DR PDB: 1SVN; 14-OCT-96.
 DR MEROPS: S08.103.1.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE ASP; 1.
 DR PROSITE: PS00137; SUBTILASE H-S; 1.
 DR PROSITE: PS00138; SUBTILASE SR; 1.
 KW Hydrolyase; Sporulation; Serine protease; Metal-binding;
 KM Calcium-binding; 3d-structure;
 FT ACT_SITE 32 32 CHANGE RELAY SYSTEM.
 FT ACT_SITE 62 62 CHANGE RELAY SYSTEM.
 FT ACT_SITE 215 215 CHANGE RELAY SYSTEM.
 FT METAL 2 2 CALCIUM 1.
 FT METAL 40 40 CALCIUM 1.
 FT METAL 73 73 CALCIUM 1.
 FT METAL 75 75 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 77 77 CALCIUM 1.
 FT METAL 79 79 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 163 163 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 165 165 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 168 168 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT STRAND 2 2
 FT HELIX 6 6
 FT TURN 11 12
 FT HELIX 13 18
 FT TURN 19 20

FT TURN 24 25
 FT STRAND 27 32
 FT TURN 39 40
 FT STRAND 43 48
 FT TURN 51 52
 FT HELIX 62 71
 FT STRAND 78 78
 FT TURN 84 85
 FT STRAND 87 92
 FT STRAND 94 94
 FT TURN 96 97
 FT STRAND 100 100
 FT HELIX 102 114
 FT TURN 115 116
 FT STRAND 119 122
 FT STRAND 126 126
 FT HELIX 131 142
 FT TURN 143 144
 FT STRAND 146 150
 FT STRAND 161 161
 FT TURN 162 164
 FT TURN 166 167
 FT STRAND 168 174
 FT TURN 176 177
 FT STRAND 180 180
 FT TURN 182 183
 FT TURN 188 189
 FT STRAND 190 195
 FT STRAND 199 203
 FT TURN 204 206
 FT STRAND 207 211
 FT HELIX 214 231
 FT TURN 233 234
 FT HELIX 237 246
 FT TURN 247 247
 FT STRAND 249 249
 FT HELIX 254 257
 FT TURN 258 259
 FT STRAND 261 261
 FT HELIX 264 267
 FT TURN 268 269
 SQ SEQUENCE 269 AA; 26698 MW; 4D89F8778999BFD CRC64;
 Query Match 11.0%; Score 246; DB 1; Length 269;
 Best Local Similarity 31.4%; Pred. No. 2,28-09;
 Matches 92; Conservative 30; Mismatches 89; Indels 82; Gaps 14;
 QY 8 VKADYAGSSYGLYQGQIVAVADTGLDTRNDSSMEAFKRTALYALGRTNANDTNG 67
 DB 11 VOAPAAHNR-GLTSGGVAVAVLDTGIST-----HPDLNIRGASFPSP--STODGNG 61
 QY 68 HGTHVAGSV--LGNGATNKGAPOANLVFGSIMDSGGLGLENLQTLFSAQPSADARI 125
 DB 62 HGTHVAGTIALNNSIGVLGVAIPSAELYAVVYLAAG--SGSVSIAQGLEMGANGMHV 119
 QY 126 HTNENGA-----AVNGAYTTDSRNVDDYVRKNDMTLLFAAGNCPNGGTISAFGTA 176
 DB 120 ANLSIGSPSPATLEQAVNSA---TSRGV-----LVVAASGNNG--AGSISYPARY 165
 QY 177 KQAIVAGATE--NLRPSFGSYADNINHVAFSSSGPTKDGIRKPDVWAPGYIISASSSL 234
 DB 166 ANNAVAGTIDQNNRNASISQYAGI-----DIVAFGVNVOSTYTPG-- 205
 QY 235 APDSFMANHDSKYAYMGTSMACTPIYAGNYA-----QUREHFVKN 275
 DB 206 -----STYASLNGTSMATPHVAGAAALVKKQKPSMSNVQIRNH-LKN 246
 RESULT 12
 ID ELYA_BACAO STANDARD; PRT; 380 AA.
 AC P27693;
 DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 OS Bacillus alcalophilus.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NCBI_TaxID=1445;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PB92;
 RX MEDLINE=91282483; PubMed=2059048;
 RA van der Laan J.C., Gerritse G., Mulleners L.J.M., van der Hoek R.A.,
 Quax W.J.;
 RT "Cloning, characterization, and multiple chromosomal integration of a
 RT Bacillus alkaline protease gene.";
 RL Appl. Environ. Microbiol. 57:901-909 (1991).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RC STRAIN=PB92;
 RX MEDLINE=92390330; PubMed=1518788;
 RA van der Laan J.C., Teplovskoy A.V., Kelders H., Kalk K.H., Missel O.,
 Mulleners L.J.M., Dijkstra B.W.;
 RT "Crystal structure of the high-alkaline serine protease PB92 from
 RT Bacillus alcalophilus.";
 RL Protein Eng. 5:405-411 (1992).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
 RX MEDLINE=93078250; PubMed=1447775;
 RA Sobek H., Hecht H.-J., Ahle W., Schomburg D.;
 RT "X-ray structure determination and comparison of two crystal forms of
 RT a variant (Asn15Arg) of the alkaline protease from Bacillus
 RT alcalophilus refined at 1.85-A resolution.";
 RL J. Mol. Biol. 228:108-117 (1992).
 RN [4]
 RP STRUCTURE BY NMR OF 112-380.
 RC STRAIN=PB92;
 RX MEDLINE=92727237; PubMed=9115441;
 RA Martin J.R., Kulder F.A., Karim-Nejad Y., van der Zwan J.,
 Martini M., Schipper D., Boelens R.;
 RT "The solution structure of serine protease PB92 from Bacillus
 RT alcalophilus presents a rigid fold with a flexible substrate-binding
 RT site.";
 RL Structure 5:521-532 (1997).
 CC -1- COFACTOR: Binds 2 calcium ions per subunit.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- SIMILARITY: Belongs to peptidase family S8.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M65086; AAA22212.1; -
 DR EMBL: A13738; CAA0128.1; -
 DR PIR: A49778; A49778.
 DR PDB: 1AH2, 15-APR-98.
 DR MEROPS: S08.038; -
 DR InterPro: IPR000209; Peptidase_S8.
 DR InterPro: IPR009020; Protease_inhb.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Hydrolyase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
 KW Signal; 3D-structure.
 FT SIGNAL 1 27
 FT PROPEP 28 112 POTENTIAL.
 FT CHAIN 113 380 ALKALINE PROTEASE.
 FT ACT_SITE 143 143 CHARGE RELAY SYSTEM.
 FT ACT_SITE 173 173 CHARGE RELAY SYSTEM.

FT ACT SITE 326 326 CHARGE RELAY SYSTEM.
 FT METAL 113 113 CALCIUM 1.
 FT METAL 151 151 CALCIUM 1.
 FT METAL 184 184 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 186 186 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 188 186 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 190 190 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 274 274 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 276 276 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 279 279 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT HELIX 117 120 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT TURN 121 123
 FT HELIX 125 130
 FT TURN 131 131
 FT TURN 133 131
 FT TURN 135 136
 FT STRAND 138 142
 FT TURN 150 151
 FT STRAND 156 157
 FT HELIX 173 182
 FT STRAND 198 201
 FT TURN 207 208
 FT HELIX 213 226
 FT STRAND 230 233
 FT HELIX 242 252
 FT TURN 253 255
 FT STRAND 257 258
 FT TURN 267 268
 FT TURN 272 275
 FT STRAND 279 279
 FT STRAND 282 285
 FT STRAND 291 291
 FT TURN 299 300
 FT STRAND 303 306
 FT STRAND 310 314
 FT TURN 315 317
 FT STRAND 318 322
 FT HELIX 325 342
 FT TURN 344 345
 FT HELIX 348 358
 FT HELIX 375 380
 SQ SEQUENCE 380 AA; 38653 MW; 539EA72771B6682C CRC64;
 Query Match 11.0%; Score 246; DB 1; Length 380;
 Best Local Similarity 31.4%; Pred. No. 3.4e-09;
 Matches 92; Conservative 30; Mismatches 89; Indels 82; Gaps 14;
 QY 8 VKADVAASSYGLVGGQIVVADTGDTRGNDSSWEAFRGKTTALYALGRTNANDTNG 67
 DB 122 VQAPAAHNR-GLTSGGVKAVLDTGIST-----HPDLNRGGASFVPGEP-STQDNG 172
 QY 68 HGHVAGSV--LANGATNKGMAPQANLVFQSIMDSSGGLGFLPSNLQTLFSCQAFSAGARI 125
 DB 173 HGHVACTIAALNLSIGLVGVAPELVKVLGAGS--SGSVSIAQGLEWAGNGGMHY 230
 QY 126 HTSMKGA-----AVNAYTTDSRANDDYRKNDMTILFAAGNEGPNGTISAPGTA 176
 DB 231 ANSLGSPSPSATLEDAVNSA--TSRGV-----LVPAASGNSG--AGSISYPARY 276
 QY 177 KNAITVGATE--NLKPSFGSYADININHVAFSSRGPTKGRIRKPDVAPQTYLTSARSSL 224
 DB 277 ANMAVAGATDQNNRRASFSGYAGL-----DIVAGVNVQSTYPG-- 316
 QY 235 APDSGFVANDSKYATWGTSAATPVVAGNVA-----QUREFVKX 275
 DB 317 -----STYASLNGTSKATPVVAGAAALVKQKPSWSNVQIRNH-LKN 357
 RESULT 13
 ID ELVA_BACCS STANDARD; PRT; 380 AA.
 AC P41362;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

```

Beat Local Similarity 31.4%; Pred. No. 3,4e-03;
Matches 92; Conservative 30; Mismatches 89; Indels 82; Gaps 14;

Qy 8 VKADVAASSVGLVQGOIVAVADTGLDTGNDNSMHEAFRGKITALYALGRINVANDTNG 67
Db 122 VQAPAAHNR-GLTSGSVKVAVLDLTGIST-----HDLNIRGASVPPGP-STQGGG 172
Qy 68 HGHVAVSV--LNGGANKKMAFPANLVFQSIINDSGGLGLPSNLQTLTSQAFSAGARI 125
Db 173 HGHVAVDTIALNNISICVLVAAPSAELVAVKVLGASG--SGSVSSIIAQLEWAGNNQW 230
Qy 126 HTNSMGA-----AVNGAVYTTDSNVDVYRKDMTLTFLAAGNGPFGGTISAPGTA 176
Db 231 ANLSLGSPPSALTLEQAVNSA---TSGGV-----LVMASSNGG--AGSISTYARY 276
Qy 177 KNAITVGATE--NLRPSFGSYADNINHNVAQFSRSGPTKDGRIKPDYMAPGTYLISASSTL 234
Db 277 ANMAAVATQDQNNNRASFQSYGAGL-----DIVAPVNVQSTPYG- 316
Qy 235 APDSFMANHDSKAYVMGTSMATPIVAGNV-----QREHEVKN 275
Db 317 -----STYASLNGTSMATPIVAGNVAAALVKQNPSPSNVQIDNR-LKN 357

RESULT 14
THES_BACSP STANDARD; PRT; 401 AA.
ID AC 045670;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Thermophilic serine proteinase precursor (EC 3.4.21.-) (Ak.1
DE protease).
OS Bacillus sp. (strain AK1).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
CX NCBI_TaxID=1409;
[1]
RN R1
RP SEQUENCE FROM N.A.
RA MEDLINE=95085262; Pubmed=7993087;
RA Maciver B., McHale R.H., Saul D.J., Bergquist P.L.;
RT "Cloning and sequencing of a serine proteinase gene from a
RT thermophilic Bacillus species and its expression in Escherichia
RT coli".
RL Appl. Environ. Microbiol. 60:3981-3988(1994).
[2]
RN R1
RP X-RAY CRYSTALLOGRAPHY (1.80 ANGSTROMS) OF 122-401.
RA MEDLINE=20057863; Pubmed=105889904;
RA Smith C.A., Toogood H.S., Baker H.M., Daniel R.M., Baker E.N.;
RT "Calcium-mediated thermostability in the subtilisin superfamily: the
RT crystal structure of Bacillus Ak.1 protease at 1.8-A resolution.".
RL J. Mol. Biol. 294:1027-1040(1999).
CC -1- COFACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: Has a pH optimum of 8.5, a temperature optimum of
CC 75 degrees Celsius.
CC -1- SIMILARITY: Belongs to peptidase family S6.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
Cc DR EMBL; L29506; AAA63688.1; -.
Cc DR PIR; I39974; I39974.
Cc DR PDB; 1DBI; 18-NOV-99.
Cc DR MEROPS; S08.009.-.
Cc DR InterPro; IPR000209; Peptidase_S8.
Cc DR InterPro; IPR000202; Protease_fimh.
Cc DR Pfam; PF00082; Peptidase_S8; 1.
Cc DR PRINTS; PR00723; SUBTILISIN.

```

DR	PROSITE; PS00136;	SUBTILASE_ASF; 1.
DR	PROSITE; PS00137;	SUBTILASE_HIS; 1.
DR	PROSITE; PS00138;	SUBTILASE_SER; 1.
KW	Hydrolase; Serine protease; Zymogen; Metal-binding; Calcium-binding	
KW	Signal; 3D-structure.	
FT	SIGNAL	POTENTIAL.
FT	PROPEP	1
FT	CHAIN	25
FT	ACG_SITE	122
FT	ACG_SITE	160
FT	ACG_SITE	193
FT	ACG_SITE	347
FT	METAL	126
FT	METAL	126
FT	METAL	168
FT	METAL	169
FT	METAL	171
FT	METAL	171
FT	METAL	179
FT	METAL	184
FT	METAL	186
FT	METAL	204
FT	METAL	204
FT	METAL	204
FT	METAL	207
FT	METAL	209
FT	METAL	211
FT	METAL	297
FT	METAL	300
FT	METAL	323
FT	DISULFID	258
FT	TURN	127
FT	HELIX	129
FT	TURN	135
FT	HELIX	137
FT	TURN	140
FT	HELIX	141
FT	TURN	145
FT	TURN	152
FT	STRAND	155
FT	TURN	165
FT	TURN	168
FT	STRAND	174
FT	TURN	180
FT	STRAND	184
FT	HELIX	193
FT	STRAND	219
FT	TURN	228
FT	HELIX	234
FT	TURN	247
FT	STRAND	251
FT	HELIX	263
FT	TURN	275
FT	STRAND	278
FT	STRAND	285
FT	TURN	298
FT	STRAND	301
FT	TURN	308
FT	STRAND	312
FT	TURN	314
FT	STRAND	316
FT	TURN	320
FT	STRAND	324
FT	STRAND	331
FT	TURN	336
FT	STRAND	339
FT	HELIX	346
FT	TURN	363
FT	HELIX	367
FT	TURN	377
FT	STRAND	379
FT	TURN	382
FT	STRAND	384
FT	TURN	385
FT	STRAND	388
FT	STRAND	391
FT	HELIX	395

FT	THEN	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908
----	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

```
FT METAL 75 75 CALC1UM 1. (VIA CARBONYL OXYGEN) .
FT METAL 77 77
FT METAL 79 79 CALC1UM 1 (VIA CARBONYL OXYGEN) .
FT METAL 163 163 CALC1UM 2 (VIA CARBONYL OXYGEN) .
FT METAL 165 165 CALC1UM 2 (VIA CARBONYL OXYGEN) .
FT METAL 168 168 CALC1UM 2 (VIA CARBONYL OXYGEN) .
FT HELIX 6 10
FT TURN 11 12
FT HELIX 13 18
FT TURN 19 20
FT TURN 24 25
FT STRAND 27 32
FT TURN 39 40
FT STRAND 43 48
FT TURN 51 52
FT HELIX 62 71
FT TURN 84 85
FT STRAND 87 92
FT TURN 96 97
FT HELIX 102 114
FT TURN 115 116
FT STRAND 119 122
FT STRAND 126 126
FT HELIX 131 142
FT TURN 143 144
FT STRAND 146 150
FT STRAND 161 161
FT TURN 162 164
FT TURN 166 167
FT STRAND 169 174
FT TURN 176 177
FT STRAND 180 180
FT TURN 182 183
FT TURN 188 189
FT STRAND 192 195
FT STRAND 199 203
FT TURN 204 206
FT STRAND 207 211
FT HELIX 214 231
FT TURN 233 234
FT HELIX 237 246
FT TURN 247 247
FT STRAND 249 249
FT HELIX 254 257
FT TURN 258 259
FT STRAND 261 261
FT HELIX 264 267
FT TURN 268 269
SQ SEQUENCE 269 AA; 26723 MM; 7A03C86D534A1D07 CRC64;
```

```
Query Match 10.9%; Score 245; DB 1; Length 269;
Best Local Similarity 31.4%; Pred. No. 2.5e-09;
Matches 92; Conservative 30; Mismatches 89; Indels 82; Gaps 14;
```

```
QY 8 VKADVAQSSGYGQGIIVADTGLDGRNDSMEAFRGKITATLYALGRTNNANDTG 67
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 11 VQAPAHNR-GLTSGYKVALDTGIST-----HPDLNIRGASFPGEF-STDGG 61
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 68 HGTIVAGSV--LGNATNKQAPQANLVFQSIMDSGGLGGLPSNLQTLFQAFAAGAR 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 HGTIVAGTIALNNSIGLVAFSAELVAVVLGASG--SGSVSIAQGLEWAGNGMHV 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 126 HTNMGWGA-----AVNGAYITDSRNVDYVAKNDMTILFAAGNCGPNGGTISAPGTA 176
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 ANLSLGSFSPATLEQAVNSA--TSRGV-----LVVAASGNSG--AGSISTPYAR 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 177 KNATTVGATE--NURPSFGSYADNINVAQFSSRGPTKGRKIPDVWAPGTIYLSARSS 234
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 166 ANAAVAGATDQNNRRASFQYAGL-----DIVAPGVNVQSTYPG- 205
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 235 APDSSFWANHDSKYAMGCTSMATPIVAGNV-----QLREHFVK 275
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 206 -----STVASLNGISMAIPHVAGVALVKKQNPMSNVQIRNH-LKN 246
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Search completed: March 31, 2004, 16:05:27
Job time : 9.51154 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 32.55 Seconds
(Without alignments)
4206.909 Million cell updates/sec

Title: US-09-985-689A-2

Perfect score: 2245
Sequence: 1 NDVARGIVKADVAQSSXGIV.....EVQAVNPVGPQNSLAIYN 434

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriapi:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2237	99.6	639	2 Q9AOR3	Q9AGR3 bacillus sp
2	2191	97.6	640	2 Q93UV9	Q93UV9 bacillus sp
3	2130	94.9	434	2 Q9AOR0	Q9AGR0 bacillus sp
4	2010.5	89.6	433	2 Q9AOR1	Q9AGR1 bacillus sp
5	2006.5	89.4	433	2 Q9AOR4	Q9AGR4 bacillus sp
6	1999.5	89.1	433	2 Q9AOR2	Q9AGR2 bacillus sp
7	509.5	22.7	1925	5 Q8TSM1	Q8TSM1 dictyosteli
8	443	19.7	1702	5 Q9GTN7	Q9GTN7 dictyosteli
9	423	18.8	654	17 Q8UC09	Q8UC09 pyrococcus
10	418	18.6	561	16 Q8RB02	Q8RB02 thermococcus
11	407.5	18.2	1239	16 Q9FB24	Q9FB24 streptomyc
12	387	17.2	430	16 Q8ENV1	Q8ENV1 oceanobacil
13	385	17.1	1253	16 Q9FC06	Q9FC06 streptomyc
14	365	16.3	1102	2 P95684	P95684 streptomyc
15	360.5	16.1	1208	16 Q82B14	Q82B14 streptomyc
16	357.5	15.9	444	16 Q9KBU7	Q9KBU7 bacillus ha

17	354.5	15.8	1105	2 Q8KX6	Q8KX6 streptomyc
18	351	15.6	1245	16 Q8RL54	Q8RL54 streptomyc
19	345	15.4	1139	16 Q82I39	Q82I39 streptomyc
20	343	15.3	1237	2 Q8GGT4	Q8GGT4 streptomyc
21	340	15.1	412	2 Q9AER6	Q9AER6 thermococcus
22	339	15.1	412	16 Q8RC68	Q8RC68 oceanobacil
23	331.5	14.8	435	16 Q8EMJ3	Q8EMJ3 oceanobacil
24	329.5	14.7	1220	16 Q9L0A0	Q9L0A0 streptomyc
25	326	14.5	442	16 Q3I788	Q3I788 bacillus su
26	323.5	14.4	891	10 Q93635	Q93635 thermococcus
27	315.5	14.1	824	2 Q45464	Q45464 bacillus sp
28	306.5	13.7	1398	1 Q9P9L1	Q9P9L1 pyrococcus
29	301.5	13.4	431	2 Q9S3I6	Q9S3I6 bacillus sp
30	289	12.9	1135	1 Q9P9D1	Q9P9D1 uncultured
31	287.5	12.8	434	2 Q54327	Q54327 bacillus sp
32	283	12.6	419	2 Q45681	Q45681 bacillus su
33	280.5	12.5	586	16 Q8PAL8	Q8PAL8 xanthomonas
34	279	12.4	799	16 Q9KEM1	Q9KEM1 bacillus ha
35	275.5	12.3	959	16 Q8PMS7	Q8PMS7 xanthomonas
36	274.5	12.2	1345	1 Q54437	Q54437 staphylothe
37	274	12.2	1098	16 Q9L1Z8	Q9L1Z8 streptomyc
38	272.5	12.1	621	2 Q9F486	Q9F486 alteromonas
39	272.5	12.1	621	2 Q53401	Q53401 alteromonas
40	270	12.0	715	2 P70765	P70765 alteromonas
41	270	12.0	1101	16 Q82CFO	Q82CFO streptomyc
42	269.5	12.0	568	16 Q8PMC0	Q8PMC0 xanthomonas
43	269.5	12.0	1571	2 Q8GCM3	Q8GCM3 streptococc
44	268.5	12.0	629	2 Q93RG8	Q93RG8 alteromonas
45	265.5	11.8	403	2 Q45463	Q45463 bacillus sp

ALIGNMENTS

RESULT 1

Q9AOR3 PRELIMINARY; PRT; 639 AA.
ID Q9AOR3
AC Q9AOR3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease.
GN PROA.
OS Bacillus sp. 9860.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_taxid=133778;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9860;
RX MEDLINE=20568675; PubMed=1118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.,
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
DR EMBL; AB046403; BA321266.2; -
DR HSSP; P00782; 1SUP.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002029; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_SER; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
SQ SEQUENCE 639 AA; 68185 MW; 316AR6PDB24FF54 CRC64;
Query Match 99.6%; Score 2237; DB 2; Length 639;
Best Local Similarity 99.8%; Pred. No. 5.2e-11;
Matches 433; Conservative 0; Mismatches 1; Indels 0; Gaps 0;


```

QY 1 NDVARGIVKADVAQSSYGLGQGIYAVADTGLDGTGRNDSMHEARFGKITLALYALGRN 60
DB 206 NDVARGIVKADVAQSSYGLGQGIYAVADTGLDGTGRNDSMHEARFGKITLALYALGRN 265
QY 61 NANDTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSGAQS 120
DB 266 NANDTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSGAQS 325
QY 121 AGARHTNSWGAANNGAYTTDSRVDDYRKNDMTLLFAAGNGPRNGGTTISAGTAKNAI 180
DB 326 AGARHTNSWGAANNGAYTTDSRVDDYRKNDMTLLFAAGNGPRNGGTTISAGTAKNAI 385
QY 181 TVGATEMLRPSFGSYADININHVAFSSRGPTKDGRIKPDVMAFGTYILSARSSLAAPDSF 240
DB 386 TVGATEMLRPSFGSYADININHVAFSSRGPTKDGRIKPDVMAFGTYILSARSSLAAPDSF 445
QY 241 WANHDSKYAMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKAAIAGAADVGLGY 300
DB 446 WANHDSKYAMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKAAIAGAADVGLGY 505
QY 301 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYTFATAKPKLSIWSMDAPASTTA 360
DB 506 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYTFATAKPKLSIWSMDAPASTTA 565
QY 361 SVTLVNDLDVITAPNGTRVYVNDPSAFPDMNDGNNENVFINSPOGTYTIEVOAYN 420
DB 566 SVTLVNDLDVITAPNGTRVYVNDPSAFPDMNDGNNENVFINSPOGTYTIEVOAYN 625
QY 421 VPVGPQNFSLAIYN 434
DB 626 VPVGPQNFSLAIYN 639

RESULT 2
Q93UV9 PRELIMINARY; PRT; 640 AA.
AC Q93UV9;
DT 01-JUN-2001 (TrEMBLrel. 19, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease.
GN PROF.
OS Bacillus sp. KSM-KP43.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxID=109322;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KP43;
RA Itoh S., Saeki K.;
RT "new protease.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB051423; BAB55674.2;
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002039; Peptidase_S8.
DR Pfam; PF04151; PPC; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR PROSITE; PS00139; SUBTILASE_SER; 1.
SQ SEQUENCE 640 AA; 67991 MW; 48BAF77E9D592C15 CRC64;

Query Match 97.6%; Score 2191; DB 2; Length 640;
Best Local Similarity 96.5%; Pred. No. 2.3e-118;
Matches 419; Conservative 13; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 NDVARGIVKADVAQSSYGLGQGIYAVADTGLDGTGRNDSMHEARFGKITLALYALGRN 60
DB 207 NDVARGIVKADVAQSSYGLGQGIYAVADTGLDGTGRNDSMHEARFGKITLALYALGRN 266
QY 61 NANDTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSGAQS 120
DB 61 NANDTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSGAQS 120

```

```

DB 267 NANDTNGHGHVAGSVLGNATNKGAPOANLVFQSIIMDSGGGLGGLPSNLQTLFSGAQS 326
QY 121 AGARHTNSWGAANNGAYTTDSRVDDYRKNDMTLLFAAGNGPRNGGTTISAGTAKNAI 180
DB 327 AGARHTNSWGAANNGAYTTDSRVDDYRKNDMTLLFAAGNGPRNGGTTISAGTAKNAI 386
QY 181 TVGATEMLRPSFGSYADININHVAFSSRGPTKDGRIKPDVMAFGTYILSARSSLAAPDSF 240
DB 387 TVGATEMLRPSFGSYADININHVAFSSRGPTKDGRIKPDVMAFGTYILSARSSLAAPDSF 446
QY 241 WANHDSKYAMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKAAIAGAADVGLGY 300
DB 447 WANHDSKYAMGTSVATPIVAGNVAQLREHFVKNRGITPKPSLLKAAIAGAADVGLGY 506
QY 301 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYTFATAKPKLSIWSMDAPASTTA 360
DB 507 PNGNGGGRVTLDKSLNVAAYVNESSALSTSOKATYTFATAKPKLSIWSMDAPASTTA 566
QY 361 SVTLVNDLDVITAPNGTRVYVNDPSAFPDMNDGNNENVFINSPOGTYTIEVOAYN 420
DB 567 SVTLVNDLDVITAPNGTRVYVNDPSAFPDMNDGNNENVFINSPOGTYTIEVOAYN 626
QY 421 VPVGPQNFSLAIYN 434
DB 627 VPVGPQNFSLAIYN 640

RESULT 3
Q9AOR0 PRELIMINARY; PRT; 434 AA.
AC Q9AOR0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Protease (Fragment).
GN PROF.
OS Bacillus sp. NVL.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxID=133781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NVL;
RA MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; AB046406; BAB21269.1;
DR HSSP; P00782; ISUP.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002039; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR PROSITE; PS00139; SUBTILASE_SER; 1.
DR Hydrolase; Hydrolase; Serine protease.
FT NON_TER 1
FT NON_TER 434
SQ SEQUENCE 434 AA; 45294 MW; 83517EDDB74125D2 CRC64;

Query Match 94.9%; Score 2130; DB 2; Length 434;
Best Local Similarity 93.4%; Pred. No. 4.6e-115;
Matches 401; Conservative 24; Mismatches 9; Indels 0; Gaps 0;

```

```

QY 1 NDVARGIVKADVAQSSYGLGQGIYAVADTGLDGTGRNDSMHEARFGKITLALYALGRN 60
DB 1 NDVARGIVKADVAQSSYGLGQGIYAVADTGLDGTGRNDSMHEARFGKITLALYALGRN 60

```

Db 1 NDVARGIVKADVAOSSYGLYGQGVVAVADTGLDTRGNDSSMHEARFGKITAYALGRTN 60
 QY 61 NNDNTNGHGTTHVAGSYLNGATNKGAPOANLVFQSIIMDSGGGLGSLPSNLTQLFSGQAFS 120
 Db 61 NNDNTNGHGTTHVAGSYLNGATNKGAPOANLVFQSIIMDSGGGLGSLPSNLTQLFSGQAFS 120
 QY 121 AGARITHNSWGAAYNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPGGTTISAPGTAKNAI 180
 Db 121 AGARITHNSWGAAYNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPGGTTISAPGTAKNAI 180
 QY 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTXDKGRKPDVMAFGTYILSARSLAPDSSF 240
 Db 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTXDKGRKPDVMAFGTYILSARSLAPDSSF 240
 QY 241 WANHDSKYAMGTSMTATPIVAGNVAQLRHEHFKNGGITPKPSLLKALIAAGADVGLGY 300
 Db 241 WANHDSKYAMGTSMTATPIVAGNVAQLRHEHFKNGGITPKPSLLKALIAAGADVGLGY 300
 QY 301 PNGNQGWGRVTLDKSLNVAAYVNESSALSTSQKATYTFRTAGKPLKISLWSDAPASTTA 360
 Db 301 PNGNQGWGRVTLDKSLNVAAYVNESSALSTSQKATYTFRTAGKPLKISLWSDAPASTTA 360
 QY 361 SYTLVNDLDLVITAPNGTRYVGNDFSAFPDNNMDGNNVENVFINSPOGTYTIEVOAYN 420
 Db 361 SYTLVNDLDLVITAPNGTRYVGNDFSAFPDNNMDGNNVENVFINSPOGTYTIEVOAYN 420
 QY 421 VPVGQPNFSLATYN 434
 Db 421 VPVGQPNFSLATYN 434

RESULT 4

Q9AQR1 ID Q9AQR1 PRELIMINARY; PRT; 433 AA.
 AC Q9AQR1;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROD.
 OS Bacillus sp. SD521.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NC NCB1_TaxID=133780;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SD521;
 RX MEDLINE=2056675; PubMed=11118284;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H., Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from alkaliphilic Bacillus spp.: enzymatic properties, sequences, and evolutionary relationships.";
 RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046405; BAB21268.1; -.
 DR HSSP; Q45670; IDB1.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR007280; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 FT NON_TER 1 433
 SQ SEQUENCE 433 AA; 45576 MW; 98A2DF18F860DC CRC64;

Query Match 89.6%; Score 2010.5; DB 2; Length 433;
 Best Local Similarity 88.5%; Pred. No. 3.5e-108;

Matches 384; Conservative 26; Mismatches 23; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAOSSYGLYGQGVVAVADTGLDTRGNDSSMHEARFGKITAYALGRTN 60
 Db 1 NDVARGIVKADVAOSSYGLYGQGVVAVADTGLDTRGNDSSMHEARFGKITAYALGRTN 60
 QY 61 NNDNTNGHGTTHVAGSYLNGATNKGAPOANLVFQSIIMDSGGGLGSLPSNLTQLFSGQAFS 120
 Db 61 NNDNTNGHGTTHVAGSYLNGATNKGAPOANLVFQSIIMDSGGGLGSLPSNLTQLFSGQAFS 120
 QY 121 AGARITHNSWGAAYNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPGGTTISAPGTAKNAI 180
 Db 121 AGARITHNSWGAAYNGAYTTDSRNVDYVRKNDMTLLFAAGNEGPGGTTISAPGTAKNAI 180
 QY 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTXDKGRKPDVMAFGTYILSARSLAPDSSF 240
 Db 181 TVGATENLRPSFGSYADNINHVAFQSSRGPTXDKGRKPDVMAFGTYILSARSLAPDSSF 240
 QY 241 WANHDSKYAMGTSMTATPIVAGNVAQLRHEHFKNGGITPKPSLLKALIAAGADVGLGY 300
 Db 241 WANHDSKYAMGTSMTATPIVAGNVAQLRHEHFKNGGITPKPSLLKALIAAGADVGLGY 300
 QY 301 PNGNQGWGRVTLDKSLNVAAYVNESSALSTSQKATYTFRTAGKPLKISLWSDAPASTTA 360
 Db 301 PNGNQGWGRVTLDKSLNVAAYVNESSALSTSQKATYTFRTAGKPLKISLWSDAPASTTA 360
 QY 361 SYTLVNDLDLVITAPNGTRYVGNDFSAFPDNNMDGNNVENVFINSPOGTYTIEVOAYN 420
 Db 361 SYTLVNDLDLVITAPNGTRYVGNDFSAFPDNNMDGNNVENVFINSPOGTYTIEVOAYN 420
 QY 421 VPVGQPNFSLATYN 434
 Db 421 VPVGQPNFSLATYN 434

RESULT 5

Q9AQR4 ID Q9AQR4 PRELIMINARY; PRT; 433 AA.
 AC Q9AQR4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROD.
 OS Bacillus sp. D6.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 NC NCB1_TaxID=127889;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=D6;
 RX MEDLINE=2056675; PubMed=11118284;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H., Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from alkaliphilic Bacillus spp.: enzymatic properties, sequences, and evolutionary relationships.";
 RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046402; BAB21265.1; -.
 DR HSSP; Q45670; IDB1.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR007280; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolase; Protease; Serine protease.
 FT NON_TER 1 433
 SQ SEQUENCE 433 AA; 45576 MW; 98A2DF18F860DC CRC64;

Query Match 89.6%; Score 2010.5; DB 2; Length 433;
 Best Local Similarity 88.5%; Pred. No. 3.5e-108;

SQ SEQUENCE 433 AA; 45636 MW; 52087E0A2516107F CRC64;
 Query Match 89.4%; Score 2006.5; DB 2; Length 433;
 Best Local Similarity 88.2%; Pred. No. 6e-108;
 Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGLDTRGRNDSVHEAFRGKITLALYALGRTN 60
 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGLDTRGRNDSVHEAFRGKITLALYALGRTN 60
 DB 61 NNDPARGHGHVAGSVLGN-ALNKGMAPOANLVFOSIMSSGGLGGLPSNLMTLTFEQAWN 119
 61 NNDPARGHGHVAGSVLGN-ALNKGMAPOANLVFOSIMSSGGLGGLPSNLMTLTFEQAWN 119
 QY 121 AGARHTNSGAAVNGAYTTDSRNVDYRKNDMTLFAAGNEGPNGGTISAPGTAKNAI 180
 121 AGARHTNSGAAVNGAYTTDSRNVDYRKNDMTLFAAGNEGPNGGTISAPGTAKNAI 180
 DB 120 AGARHTNSGAAVNGAYTTDSRNVDYRKNDMTLFAAGNEGPNGGTISAPGTAKNAI 179
 120 AGARHTNSGAAVNGAYTTDSRNVDYRKNDMTLFAAGNEGPNGGTISAPGTAKNAI 179
 QY 181 TVGATENLRPSFGSVADNINHVAQFSSRGPTKGRPKPVMAAGTYILSARSLADSSSF 240
 181 TVGATENLRPSFGSVADNINHVAQFSSRGPTKGRPKPVMAAGTYILSARSLADSSSF 240
 DB 180 TVGATENLRPSFGSVADNINHVAQFSSRGPTKGRPKPVMAAGTYILSARSLADSSSF 239
 180 TVGATENLRPSFGSVADNINHVAQFSSRGPTKGRPKPVMAAGTYILSARSLADSSSF 239
 QY 241 WANHSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVLGY 300
 241 WANHSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVLGY 300
 DB 240 WANHSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVLGY 299
 240 WANHSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVLGY 299
 QY 301 PNGNGGGRVTLDKSLNVAVNESALSTSOXATYFTATAGKPKLSLWSDAPASTTA 360
 301 PNGNGGGRVTLDKSLNVAVNESALSTSOXATYFTATAGKPKLSLWSDAPASTTA 360
 DB 300 PNGNGGGRVTLDKSLNVAVNESALSTSOXATYFTATAGKPKLSLWSDAPASTTA 359
 300 PNGNGGGRVTLDKSLNVAVNESALSTSOXATYFTATAGKPKLSLWSDAPASTTA 359
 QY 361 SVTLVNDLDLVITAPNGTRYVNDPSAPPDNNMDGNNVENVFINSPOSGTYITIEOVAYN 420
 361 SVTLVNDLDLVITAPNGTRYVNDPSAPPDNNMDGNNVENVFINSPOSGTYITIEOVAYN 420
 DB 360 SVTLVNDLDLVITAPNGTRYVNDPSAPPDNNMDGNNVENVFINSPOSGTYITIEOVAYN 419
 360 SVTLVNDLDLVITAPNGTRYVNDPSAPPDNNMDGNNVENVFINSPOSGTYITIEOVAYN 419
 QY 421 VPVGQPONFSLAIYN 434
 421 VPVGQPONFSLAIYN 434
 DB 420 VPVGQPONFSLAIYN 433
 420 VPVGQPONFSLAIYN 433

RESULT 6
 Q9AQR2 PRELIMINARY; PRT; 433 AA.
 ID Q9AQR2
 AC Q9AQR2
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROC.
 OS Bacillus sp. Y.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=133779;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y.
 RA MEDLINE=20566675; PubMed=1118284;
 RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 RA Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
 RT evolutionary relationships.";
 RT Biochem Biophys Res Commun 279:313-319 (2000).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046404; BAB21267.1; -;
 DR HSSP; Q45670; 1DBI.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR007280; PPC.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PRO0723; SUBTILISIN.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Hydrolyase; Protease; Serine protease.
 FT NON_TER 1
 FT NON_TER 433
 FT NON_TER 433
 SQ SEQUENCE 433 AA; 45587 MW; B81291A803C775AE CRC64;
 Query Match 89.1%; Score 1999.5; DB 2; Length 433;
 Best Local Similarity 88.0%; Pred. No. 1.5e-107;
 Matches 382; Conservative 27; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGLDTRGRNDSVHEAFRGKITLALYALGRTN 60
 1 NDVARGIVKADVAQSSSYGLYGQGIIVAVADTGLDTRGRNDSVHEAFRGKITLALYALGRTN 60
 DB 61 NNDPARGHGHVAGSVLGN-ALNKGMAPOANLVFOSIMSSGGLGGLPSNLMTLTFEQAWN 119
 61 NNDPARGHGHVAGSVLGN-ALNKGMAPOANLVFOSIMSSGGLGGLPSNLMTLTFEQAWN 119
 QY 121 AGARHTNSGAAVNGAYTTDSRNVDYRKNDMTLFAAGNEGPNGGTISAPGTAKNAI 180
 121 AGARHTNSGAAVNGAYTTDSRNVDYRKNDMTLFAAGNEGPNGGTISAPGTAKNAI 180
 DB 120 AGARHTNSGAAVNGAYTTDSRNVDYRKNDMTLFAAGNEGPNGGTISAPGTAKNAI 179
 120 AGARHTNSGAAVNGAYTTDSRNVDYRKNDMTLFAAGNEGPNGGTISAPGTAKNAI 179
 QY 181 TVGATENLRPSFGSVADNINHVAQFSSRGPTKGRPKPVMAAGTYILSARSLADSSSF 240
 181 TVGATENLRPSFGSVADNINHVAQFSSRGPTKGRPKPVMAAGTYILSARSLADSSSF 240
 DB 180 TVGATENLRPSFGSVADNINHVAQFSSRGPTKGRPKPVMAAGTYILSARSLADSSSF 239
 180 TVGATENLRPSFGSVADNINHVAQFSSRGPTKGRPKPVMAAGTYILSARSLADSSSF 239
 QY 241 WANHSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVLGY 300
 241 WANHSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVLGY 300
 DB 240 WANHSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVLGY 299
 240 WANHSKYAVMGSTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALTAGADVLGY 299
 QY 301 PNGNGGGRVTLDKSLNVAVNESALSTSOXATYFTATAGKPKLSLWSDAPASTTA 360
 301 PNGNGGGRVTLDKSLNVAVNESALSTSOXATYFTATAGKPKLSLWSDAPASTTA 360
 DB 300 PNGNGGGRVTLDKSLNVAVNESALSTSOXATYFTATAGKPKLSLWSDAPASTTA 359
 300 PNGNGGGRVTLDKSLNVAVNESALSTSOXATYFTATAGKPKLSLWSDAPASTTA 359
 QY 361 SVTLVNDLDLVITAPNGTRYVNDPSAPPDNNMDGNNVENVFINSPOSGTYITIEOVAYN 420
 361 SVTLVNDLDLVITAPNGTRYVNDPSAPPDNNMDGNNVENVFINSPOSGTYITIEOVAYN 420
 DB 360 SVTLVNDLDLVITAPNGTRYVNDPSAPPDNNMDGNNVENVFINSPOSGTYITIEOVAYN 419
 360 SVTLVNDLDLVITAPNGTRYVNDPSAPPDNNMDGNNVENVFINSPOSGTYITIEOVAYN 419
 QY 421 VPVGQPONFSLAIYN 434
 421 VPVGQPONFSLAIYN 434
 DB 420 VPVGQPONFSLAIYN 433
 420 VPVGQPONFSLAIYN 433

RESULT 7
 O8T9W1 PRELIMINARY; PRT; 1825 AA.
 ID O8T9W1
 AC O8T9W1
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Serine protease/ABC transporter TagD.
 GN TAGD.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
 OC NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Ax4;
 RA Anjard C., Loomis W.F.;
 RT "Evolution of the ABC transporters of Dictyostelium.";
 RT submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL; AF46309; AAL74253.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
 DR GO; GO:0000166; F:nucleotide binding; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR001140; ABC_TM_transp.

```

DR InterPro: IPR003439; ABC transporter.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00664; ABC_membrane_1.
DR Pfam: PF00005; ABC_tran; 1.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR ProDom: PD000006; ABC transporter; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR ATP-binding; Protease; Transport.
KW SEQUENCE 1825 AA; 202641 MW; E28160BC78613A3B CRC64;
SQ
Query Match 22.7%; Score 509.5; DB 5; Length 1825;
Best Local Similarity 27.7%; Pred. No. 86-21;
Matches 163; Conservative 79; Mismatches 153; Indels 193; Gaps 22;
QY 19 IYGGQGIYAVADTGLDTR--NDS-----SMEAFRGKITLALYALGRTNANDTNGH 68
DB 327 LRKSGQQLSIADTGLDGHCHCFSDSNRPYPNSVNLNHRKVVITYGSL--HDNEDYVDGH 384
QY 69 GTHVGSVLGN-----ATKGMAPQANLVFQSI--WSSGGLGLPSNLQTLFSAQFS 120
DB 385 GTHVCGSAAGAPDSSLAISFSGLATDKIAFPDLASDPNNNEFVPEEDYSQLYPLYN 444
QY 121 AGARHTNSMGA---AVNGAYTDSRVNDYVRKN--DMTILFAAGNEGPNGTISAPGT 175
DB 445 AGARVHGDWGSLSIGIGYSYSDAGSIDFLYHPDILIRAAAGNEQYSLSLS-QAT 503
QY 176 AKKATVGAENTLRPF-----
DB 504 AKKAVITVGAEQTHSEYTTDALEYSNFEVASKSTNSLCQSPDDRYCTYTTAOCCTEXT 563
QY 193 -----GSYAD-----NINVAQFSRRPTDGRIRKPDVMAFGTYILSA 230
DB 564 VNGLSGCTSYIKNSYASIFSSOPPELYNNNISCFSKQPTDGLKPDIVAPGYITSA 623
QY 231 RSLA-----PDSSFWANHDSKAYNGGTSMAPIYAGVAQLREH----- 271
DB 624 RSNAGATTDOCGDGS--PNTNALLSE-----SGTSMATPLATAATTIRQTLVDGYPT 676
QY 272 --FVKRGTTPKPSLKALILAGADVGLGYP-----NGNGMGARV 310
DB 677 GSIVESNKIQPTGSLIKALMINNAQLNCTPPLSNTNPSNAVDPFAGANFVQGWGL 736
QY 311 TLDKSLNVAYNSS-----ALSTSQKATYTF----- 338
DB 737 RSEEWL---YVESSGYKPKPSRWVIGELGDKKASNMWKEYSLSTGQWVSYCFYKPPSS 793
QY 339 --ATAKRP-LKISLVNSDAPASTTASVTLVNDLVIIT-----APNSTRVVG--ND 384
DB 794 GNSGSGIPIVATLVMTDPPSYSGAKNTVNNLDLMTNTSESEFTFYNSGSSSYNGK 853
QY 385 FSAFPMNWDGRNVANVF---INSPOGCTYIEQAVNVVPGPNFS 429
DB 854 TLPLDQ---DSINNVGIIYTPINTKSIISFPFIAGTNIPIGPONFS 898
RESULT 8
ID 09GINT7 PRELIMINARY; PRT; 1702 AA.
AC 09GINT7;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE TAGA.
GN TAGA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
CX NCBI_TaxID=4689;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Good J.R., Cabral M., Kuspa A.;
RT "Taga, a putative serine protease/ABC transporter of Dictyostelium
RT that is expressed at the onset of development and is required for the
RT differentiation of a subpopulation of preprope cells."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AF263455; A011416.1; -.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO: GO:0000166; F:nucleotide binding; IEA.
DR GO: GO:0004289; F:subtilase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001140; ABC_TM_transp.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00664; ABC_membrane_1.
DR Pfam: PF00005; ABC_tran; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR ProDom: PD000006; ABC transporter; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 1.
DR ATP-binding; Transport.
KW SEQUENCE 1702 AA; 187103 MW; 4A67716303CB7131 CRC64;
SQ
Query Match 19.7%; Score 443; DB 5; Length 1702;
Best Local Similarity 25.2%; Pred. No. 56-17;
Matches 137; Conservative 97; Mismatches 169; Indels 140; Gaps 21;
QY 18 GLYGGQGIYAVADTGLDTR--NDSMEAFRGKITLALYALGRTNANDTNGHGVAG 74
DB 300 GIKDGEIVGCADTGIDINHCFFYDINPISGSTRKILIS--YSSGNDQIDRIDGHGHIYV 358
QY 75 SVLNGAKTN-----KGMAPQANLVFQSI--WSSGGLGLPSNLQTLFSAQFSAGARHTNS 128
DB 359 TIIGSTTVDPSVSEFSGGAPNSKVAFAVDLQVSGNGLSISQNLTAIYQSTYDQNAVHCD 418
QY 129 SWGAIVGAYTDSRVNDYVRKN--DMTILFAAGNEGPNG-----TISAPGTAKNAITV 182
DB 419 AMNSNIGPFYTGTEMIDRFQWHPDFIVYRSAGN--NVNPFNSIYTLQGESYSKSLVY 477
QY 183 GATENLRPSFGSYADNIN-----
DB 478 GSN--QPS--STYSSLDYWDMDFIYNSIRTSYCTGQSGIYGTCSGDPVPTQTSVDIQO 534
QY 202 -----VAQFSRRPTDGRIRKPDVMAFGTYILSA 230
DB 535 CCSNPILAKICCTETIQOQYQNTSTYSEFIPSLFSGVETSGRKLPLLAAGSRIIS 594
QY 231 RSLAPDSSFWANH-----DSKYAVNGGTSMAPIYAGVAQLRE-----HVKNR 276
DB 595 R-SLGPSSIT--NHCSPTSGIATSAIIMIEGSSQAAVAATSAVYVROYRDGYFINCK 651
QY 277 -----GITPRLSLKALILAGAA---DVG.LGYENGQGWGRVTLDSLVNVAYNSS--- 325
DB 652 VNSVGFQPSALYKATLVNTASINDSTLEY---SQGFNIDSLKILITTNMQTISLDI 708
QY 326 -----ALSTSQKATYTFIATAGKPKLISLVNSDAPASTTASVTLVNDLVIIT--P 375
DB 709 PSSIEKADPIINTGETNYSYCFSLDSKADIDITLVMTDPPASPLSTFLVNNLDLALAFV 768
QY 376 NG--TRYVNDFSAPFPNN---WDGRNVANVFINSPOGCTYIEQAVNVVPGPNFS 429
DB 769 DGEIISTYSGNS-ETIFNTSQVITFDOLNVEVIRIDAPIGSIDVKIFGINIYIPNOSIS 827
QY 430 LAI 432
DB 828 VVI 830

```

RESULT 9

Q8BU09 PRELIMINARY; PRT; 654 AA.
 ID 08BU09
 AC Q8BU09
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Alkaline serine protease.
 GN PF1670.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 CX NCBI_TaxId=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR GO: AEO10265; AAL81794.1; -
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004289; F:subtilase activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR002029; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SSR; 1.
 DR KEGG: K00001; Serine protease.
 KM SEQUENCE 654 AA; 70230 MW; 1CB145A5F505DB34 CRC64;

Query Match 18.8%; Score 423; DB 17; Length 654;
 Best Local Similarity 29.6%; Pred. No. 2e-16;

Matches 135; Conservative 61; Mismatches 146; Indels 114; Gaps 18;

QY 18 GLVGGQITVAVDTGLDTRGNDSSMHEAFRGKITLALYALGRINNAN-----DTNGHGH 71
 DB 154 GYDSSGTTIGIIDPID-----ASHEDLOGKV-----IGWDFVNGRSPYDDHGHGH 202
 QY 72 VAGSVLNGGATN-----KMAPQANLVFGSIM--DSGGGAGLPSNLQTLFSAFSAAGRI 125
 DB 203 VASIAAGTGAASNGKYGMAPGAKLAGIKVLGADSGSISITIKGYEMAVNDKDKIGKV 262
 QY 126 HTNSWGA-----AVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPGGTISAP 173
 DB 263 INSLGSSQSSDGTALDQAIVNAWDA-----GLVVVAAGSGPDKYITIGSP 310
 QY 174 GTAKNAITVGATENLRSPGSYADNINVAOFSSRGPTDGRKIPVMAFGTYILSARS 233
 DB 311 AAASKVITTVGA-----VDKIDVITFSFSSRGTDGRKLEPVAVPAGMNTIAARS 359
 QY 234 LAPDSSFWANHDSKYAMGTSMATPIVAGNVAQLREHFVKNRGITPK--PSILKAALTA 291
 DB 360 ---GTSMGQPIINDYTAAPGTSMAIPVAGIALALLQ-----AHPMTPEKXTALIE 409
 QY 292 GA-----ADVGLGVNGNGMGRTVLDKSLNAVYVNESSALSTSQA-----TYTFT 338
 DB 410 TADIVKDELADIVAGY-----GRVNAVKAIV--YDNVAKLVFTGYVANKGSQTHQFV 460
 QY 339 ATAGKPKLISLVSDAPASTTASVTLVNDLIVITAPNGTRVYAGNDFSAFDNNMGGRN 398
 DB 461 ISGASFYATLIVYDNNAN-----SDLDLVLYPNNQ-----VYSTAYYNG 501
 QY 399 VENVFINSPOSQTYITIEVQAVNPVGPQNFSLAIVN 434
 DB 502 FEKVGYNPTDGTWTKRVSVS---GSANQYDVVS 534

RESULT 10

Q8RB02

Q8RB02 PRELIMINARY; PRT; 561 AA.

Q8RB02
 ID Q8RB02
 AC Q8RB02
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Subtilisin-like serine proteases.
 GN APR22 OR TTE0824.
 OS Thermococcus kodakarensis.
 OC Bacteria; Firmicutes; Clostridia; Thermococcaceae;
 OC Thermococcaceae; Thermococcaceae; Thermococcaceae;
 OC NCBI_TaxId=119072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MB4 / JCM 11007;
 RA MEDLINE=21992816; PubMed=11997336;
 RA Bao Q., Tian Y., Li W., Lai X., Xuan Z., Hu S., Dong W., Yang J.,
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;
 RT "A complete sequence of T. tengcongensis genome."
 RL Genome Res. 12:689-700(2002).
 DR EMBL: AEO13049; AAM24081.1; -
 DR GO: GO:0005509; F:calcium ion binding; IEA.
 DR GO: GO:0008233; F:peptidase activity; IEA.
 DR GO: GO:0004289; F:subtilase activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR002029; Peptidase_S8.
 DR InterPro: IPR007280; PPC.
 DR InterPro: IPR009020; Protease_inhib.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR Pfam: PF04151; PPC; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SSR; 1.
 DR KEGG: K00001; Serine protease.
 KM SEQUENCE 561 AA; 59968 MW; BA9C5C52F7083A18 CRC64;

Query Match 18.6%; Score 418; DB 16; Length 561;
 Best Local Similarity 30.9%; Pred. No. 3.2e-16;

Matches 140; Conservative 66; Mismatches 157; Indels 90; Gaps 19;

QY 6 GIVRADVAQSSYGCGQIVAVDTGLDTRGNDSSMHEAFRGKITLALYALGRINNAN-- 63
 DB 155 GILTK---ARSDFGYTGKNTITAIIDTIGDGNHYLS-----GGKI-----IGKDPFINNK 201
 QY 64 ---DTNGHGHVAVSVLNGGATN-----KMAPQANLVFGSIMS--SGGLGGLPSNLQTL 114
 DB 202 TTPYDDNGHGHVAVSVLNGGATN-----KMAPQANLVFGSIMS--SGGLGGLPSNLQTL 114
 QY 115 FSQAPSAAGRIHTNSWGAIVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPGGTISAP 174
 DB 262 VQNDVYGIKVINSLSTGTSSTGDTSTSLAVNRAVDGIVVVAAGNSQAPAYTIGSP 321
 QY 175 TAKNAITVGATENLRSPGSYADNINVAOFSSRGPTDGRKIPVMAFGTYILSARS 233
 DB 322 AAERKAIYVAMADV-----GELGRPL-----ASFSSRGTPADRIKEDIAAPRNITAAK--- 371
 QY 235 APDSSFWANHDSKYAMGTSMATPIVAGNVAQLREHFVKNRGITPKPSILKAALTAGAA 294
 DB 372 ---ANSVNGSVYVTSIGTSMATPIVAGTVALMLN---ANPILTPDA---KNIINSTAK 419
 QY 295 DVGLGVNGNGMGRTVLDKSLNAVYVNESSALSTSQA-----TYTFT 339
 DB 420 SMGPPSKVNYDYGAGRLGCEHAKIRVAGNFRGNITDVPNHYIT--SGHPSGRYSDTWTENA 477
 QY 340 T-ACKPKLISLV--WSDAPASTTASVTLVNDLIVITAPNGTRVYAGNDFSAFDNNMG 395
 DB 478 TMTSYPIAITLILIDMANVNP-----DEDIYLYDPSGTLIKS-----TG 517
 QY 396 RNVENVFINSPOSQTYITIEVQAVNPVGPQNF 428

Db 518 TORSETTILIPOTGYVVKVYSYR---GSGNY 547

RESULT 11

Q9FEB24 PRELIMINARY; PRT; 1239 AA.

AC Q9FEB24; (T-EMBLrel. 16, Created)

DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)

DT 01-MAR-2001 (T-EMBLrel. 24, Last annotation update)

DE Putative secreted peptidase.

GN SC07188 OR SC8A11.16C.

OS Streptomyces coelicolor.

OC Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomycetes.

OC NCBI_TaxID=1902;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=A3(2);

RC Saunders D.C., Harris D.;

RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

RN [2] SEQUENCE FROM N.A.

RP STRAIN=A3(2);

RC Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;

RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

RN [3] SEQUENCE FROM N.A.

RP MEDLINE=97000351; PubMed=8841436;

RC STRAIN=A3(2);

RA Redebach M., Kleiser H.M., Denapate D., Eichner A., Cullum J., Knaebel H., Hopwood D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).

RN [4] SEQUENCE FROM N.A.

RP STRAIN=A3(2) / M145;

RC MEDLINE=2196410; PubMed=12000953;

RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L., Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleiser H., Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M., Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S., Huang C.-H., Kleiser T., Larke L., Murphy L., Oliver K., O'Neill S., Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S., Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K., Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J., Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2)."

RL Nature 417:141-147(2002).

EMBL: AL939130; CAC01588.1; .

DR HGGSP; Q98405; IMPT.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0004289; F:protease activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR00137; PA.

DR InterPro; IPR000209; Peptidase_S8.

DR Pfam; PF02225; PA.1.

DR Pfam; PF00082; Peptidase_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE_ASP; 1.

DR PROSITE; PS00137; SUBTILASE_HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.

KW Complete proteome.

SO SEQUENCE 1239 AA; 128505 MW; 8F5E9AC6E8B1260A CRC64;

Query Match 18.2%; Score 407.5; DB 16; Length 1239;

Best Local Similarity 33.3%; Pred. No. 3.7e-15;

Matches 133; Conservative 40; Mismatches 145; Indels 73; Gaps 12;

QY 8 VKADVAAOSY-----GLYGGQIVAVADTGLDTGRNDSMHEAFRGKTLALYAG 57

Db 219 VEADLADSTAGIAPAPMAGANTGQVAVLDTGVAG-----HPDLADRIAAQSFV 272

QY 58 RTNNANDTNGHTVAGSVLGNATN---KMAPQANLVFQSIMDSGGLGSPNLQT 113

Db 273 PDENTDDRDGHGTVAATIGTGAASAGKEGVAPGARLSIGKVLDNS -GRGQISWTIAA 331

QY 114 LFSQAFSAGARIHNSMGAAVNGAVTLD -SRNVDDYVRKXDMTLFLPAAGNEGPGRTISA 172

Db 332 MEMAIVEERHAKIVVNSIGSGEQSDGSDPMGRVADRLSAQGFVVAAGN -GGEAGSITGA 390

QY 173 PGTKAKITVQATENTLPSFGSVADNINHYAOFSSRGPTDGRKIPDVAPGYTILSARS 232

Db 391 PGVATSAITVGA-----VDADTLAPSSQGFVADGALKPEITAPGVGIIAA-- 437

QY 233 SLADDSFFMNHDSKAVVMGTSMAATIVAGNVAQLREHVNRKRGITPKSLKAL--- 289

Db 438 ---NSSPAGGNAGVYSLGTSNATPHVAGAAALL-----AARPDLSGSLKDV 484

QY 290 IAGADVGLGYPNQGMGRVTLDPKSLN-----VAVNNESSALSTQKATYTPATAG 342

Db 485 LASSHRTPRDARQAGSRVDVAARVAVASATVAFGSSPGVPRRLVYTTNTTGA 544

QY 343 KPLKISLVMSDA-----PASTTASVTLVND 367

Db 545 VTLELVAAATHPAGVERLSASRYTVPAGHTADVTLTID 583

RESULT 12

Q9ENV1 PRELIMINARY; PRT; 430 AA.

AC Q9ENV1; (T-EMBLrel. 23, Created)

DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)

DE Intracellular alkaline serine proteinase.

GN O82375

OS Oceanobacillus thelyensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.

OC NCBI_TaxID=182710;

RN [1] SEQUENCE FROM N.A.

RP STRAIN=HTE831 / DSM 14371 / JCM 11309;

RA MEDLINE=22220767; PubMed=12235376;

RX Takami H., Takaki Y., Uchiyama T.;

RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments.";

RT Nucleic Acids Res. 30:3927-3935(2002).

RL EMBL; AF004601; BAC14311.1; .

DR GO; GO:0004289; F:subtilase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR000209; Peptidase_S8.

DR Pfam; PF00082; Peptidase_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE_ASP; 1.

DR PROSITE; PS00137; SUBTILASE_HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.

KW Complete proteome.

SO SEQUENCE 430 AA; 45838 MW; 6D09A99B8C1E310F CRC64;

Query Match 17.2%; Score 387; DB 16; Length 430;

Best Local Similarity 35.1%; Pred. No. 1.4e-14;

Matches 118; Conservative 53; Mismatches 125; Indels 40; Gaps 16;

QY 2 DVARGIVKADVAAOSYGLYGGQIVAVADTGLDTGRNDSMHEAFRKIT--ALYALGRT 59

Db 121 DTASSINADVAKES -GLTGSGTIAVIDTGIHP-----HEDLEGRITGFADPVKQGT 172

QY 60 NNADTNGHGTHTVAGSVLGNATN---KMAPQANLVFQSIMDSGGLGSPNLQTLF 115

Db 173 EPYDD -NGHGTHTVAGSVLGNATN---KMAPQANLVFQSIMDSGGLGSPNLQTLF 229

QY 116 SQAFSAGARIHNSMGAAVNGAVTLDSTSRNVDDYVR-----KNDMTLFLPAAGNEGPGRT 169

```

Db      230 DWICIONOSKYNINILISTL-GSDATEPAEGDPVNAVETAMDGMVYCAVAGNSGDDT 288
Qy      170 ISAGTKNAITVATENLRPSFGSYADNINHVAQFSRPTDGRKPDVMAFGYIIS 229
      289 VGSFGISPKYITVADADNNTAERS---DSSVAEFSRGPITDGLTKPPLTTPGVDIYS 344
Qy      230 ARS--SLAPSSFWANHDSKYAVMGTSMAPIVAGNAQLREHFVYKNGITPKPSILKA 297
      345 LRAGSGFIDKTKNSARVGSVNYISLSTGSMATPICAGIAGQLQ---SDSSLT--PQVKE 399
Qy      288 ALIAGADVGLGYNGNGMGRTLDKSLNVAAYNE 323
      400 KLMACODLGQS-PN-VQAGAYL--NAANLININE 430

RESULT 13
Q9FC06 PRELIMINARY; PRT; 1253 AA.
AC 09FC06;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Putative secreted peptidase.
SC07176 OR SC8A11.04C.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomyces; Streptomyces.
OX NCBI_Taxid=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Saunders D.C., Harris D.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cardeno A.M., Parkhill J., Barrall B.G., Rajandream M.A.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA MEDLINE=97000351; Pubmed=8843436;
RA Redendach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
RA Khasni H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=21956410; Pubmed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Frazer A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser H., Jarke L., Murphy J., Oliver K., O'Neil S.,
RA Rabinowitz B., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrall B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
RT EMBL; A0939130; CAC01576.1; -.
DR HSSP; Q99405; IMPT.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001317; PA.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF02225; PA; 1.
DR Pfam; PF00082; Peptidase_S8; 1.

```

```

DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR Complete proteome.
SQ SEQUENCE 1253 AA; 130971 MW; AA69B417EEEDB89 CRC64;

Query Match 17.1%; Score 385; DE 16; Length 1253;
Best Local Similarity 29.4%; Pred. No. 7,4e-14;
Matches 140; Conservative 52; Mismatches 162; Indels 122; Gaps 19;

Qy      16 SYGLYGQGIIVAAVADGLDTRNDSSMHEAFRGKITALLALGRTNANDNGQTHVAGS 75
      233 SGGVTGEGVAVAVDGTGDAG-----HPDFAGRIATAFVPDQVTRDNGGTHVAST 266
Db      76 VLNGATN---KMAPOANLVFQSIMDSSG-----GLGG 106
      287 VAGTGAASGVKEKVAIPGASLHIGKVLDSGSGDSQSVTLAEMENAVDDQAKIVSLSLD 346
Qy      107 LPSNLQTLFSGAFSAGARHTNSWGAAVNGAYTTDSRNVDDVYRKNDMTILFAAGNEPN 166
      347 SPTDGTDLSEAV-----NWLSAETGA-----LFVAAAGNGEPE 380
Qy      167 GGTISAPGTAKNATTVGATENTLRPSFGSYADNINHVAQFSRGP-TRDGRIKPDVMAFGT 225
      381 AYTGTTPAADAAALTVAVNG--FGKG---VDQLDFSSRGRVGDNAVKEDLTAPGV 433
Qy      226 YILSARSLAPDSSFANHDSKYAVMGTSMAPIVAGNAQLREHFVYKNGITPKPSIL 285
      434 GVLAARSRVAREG-----EGAYQSLGTSMAPIVAGNAALAAHEPDMTG-----QRL 482
Qy      286 KAALIAADVGLGYENNGMGRTLDKSLNVAAYNESSALSTSQ-----KATYT 336
      483 KEALVGTAGTQRSP-FDAGSGRVDAAVRSTLNSGDAFQAHPYTPGQVTRSDVT 541
Db      337 FTATAGKPKLSTLYWSDA-----PASTATVTLVNDLDTVTPANG--- 377
      542 YTNCGPAVADLTLSTALEPREGFTLSAQVYVPAIGTASVGVITILD--AAEDNGAYA 599
Qy      378 TRYV--GNDS---APFDNNMDSGRNNVENVFI---NSPQSGTYIT-EVQAVNVP 422
      600 TLVASGADGAVLARTEPVGVNKEGRATLTALTAKDHHKPLSGTIVLKDVERNTAP 655
Db

RESULT 14
P95684 PRELIMINARY; PRT; 1102 AA.
AC P95684;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Subtilisin-like protease.
OS Streptomyces albobacillus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomyces; Streptomyces.
OX NCBI_Taxid=1887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S-3253;
RX MEDLINE=97144528; Pubmed=8990295;
RA Suzuki M., Teguchi S., Yamada S., Kojima S., Miura K., Momose H.;
RT "A novel member of the subtilisin-like protease family from
RT Streptomyces albobacillus.";
RL J. Bacteriol. 179:430-438(1997).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; D83672; BAAL2040.1; -.
DR HSSP; P00782; 28BT.
DR MEROPS; S08.069; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002860; GH_BNR.
DR InterPro; IPR000209; Peptidase_S8.

```

DR Pfam; PF002012; BNR. 2.
 DR Pfam; PF000823; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KM Hydrolyase; Protease; Serine protease.
 SQ SEQUENCE 1102 AA; 114128 MW; F9E4AD2590FE559E CRC64;

Query Match 16.3%; Score 365; DB 2; Length 1102;
 Best Local Similarity 31.3%; Pred. No. 8.8e-13;
 Matches 146; Conservative 56; Mismatches 177; Indels 88; Gaps 19;

QY 3 VAR-----GIKAVAGS-----SYLVGGQIVAVADTGLDTGNDSMGEAFRG 48
 DB 183 VAAVWLDGVKASLDTSVGQIGTPKMEAGYDCKGVKIVLDTGVD-----ATHPDLRG 236
 QY 49 KITALVALGRNNANDTNGHGHVAGSVLGNCA---TNKMAPQANTLVFQSIMDSGGL 104
 DB 237 QVTASKNFTGAPTGGDVVGHGTHVASLAAGTGAQSGTKGVAPGAKIINGKVLDDAG-- 294
 QY 105 GGIPLSNLQITFSQAFSAGRIHTNSGVAVNGAYTTDSKNVDYRK--NDMTILF--AA 160
 DB 295 FGDSGSLAGMEWAAAQAGDIYMSLG---GMDTPEITDPLEAAVDKLSAEKGIIFALAA 350
 QY 161 GNEGPNGTISAPGTAKNATTVGATENLRPSFSYADNINHYAQFSSRGP-TRDGRKPD 219
 DB 351 GNGSPQ--SIGSPGASDHALTVGA-----VDDCKLADPSSTPRIGDGVKPD 397
 QY 220 VMAPGYILSARSSLPADSSFWANHDSKYAYNGTSMATPIVAGVAQUREHFVNRGRT 279
 DB 398 LTPAGVDITAAKAGNDIAKEVEKEPAGYTTISGTSMATPHVAGAAALIKQHP----- 452
 QY 280 PKRSLKAAALIAAGVGLGYPNGNGMGRVTLDSLVNAYNESALS----- 328
 DB 453 WKRAELKALITATXG--KITPFEQSGRVOYDKAITGTVAEPVSLFVGQMPHAD 510
 QY 329 --TSQATYITFTTAKPKIKSLVMSD-----APAS--TTASVTLVNDLVLITAP--NGT 378
 DB 511 KPYTKLTYRNLGTEDVTLKLTSTATGPKKAPAGFFLGASTL-----TYPANGT 562
 QY 379 RYVGNDFSAFPDNNMGRNNVENVFINSPS-----GTITIEVQANV 421
 DB 563 ASVDVTAIDTRLGAVDGYSAVVATGAGQSVRTAAVAEREVESEYNV 609

RESULT 15
 Q82B14 PRELIMINARY; PRT; 1208 AA.
 AC Q82B14;
 DT 01-UTN-2003 (TEMBLrel. 24, Created)
 DT 01-UTN-2003 (TEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Putative peptidase.
 GN SAV5721.
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxId=33903;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa U., Hanamoto A., Takahashi C.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;

RX MEDLINE=22608306; PubMed=12692562;
 RA Ikeda H., Ishikawa U., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis.";
 RL Nat. Biotechnol. 21:525-531(2003).
 DR EMBL; AP005044; BAC73433.1;
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR003137; PA.
 DR InterPro; IPR002029; Peptidase_S8.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE ASP; 1.
 DR PROSITE; PS00137; SUBTILASE HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KM Complete proteome.
 SQ SEQUENCE 1208 AA; 125548 MW; E650B5E3ABE512B CRC64;

Query Match 16.1%; Score 360.5; DB 16; Length 1208;
 Best Local Similarity 31.5%; Pred. No. 1.8e-12;
 Matches 129; Conservative 53; Mismatches 161; Indels 67; Gaps 17;

QY 8 VKADVAGSS-----YGLVGGQIVAVADTGLDTGNDSMGEAFRGKITALVALG 57
 DB 186 VEADMASSNQIGTAAADAGLTGGDVAVLDTGVDT-----THPDLAGRVSRSKFT 239
 QY 58 RTNNANDTNGHGHVAGSVLGNCA---TNKMAPQANTLVFQSIMDSGGLGILPSNLQT 113
 DB 240 DGEVADRNGHGHVITSTVGGSGAASDGTERGVAIPATLAVGVLSQGA---GSESQT 235
 QY 114 LFSQAFSA---GARHTNSGGA--AVNGAYTTD--SRVVDYRKNDMTILFPAAGNEPN 166
 DB 296 IAGMEMARVRRAIYMSLSGSTEASDG---TTPMAEAVDTLSEETGALFVVAAGNTGA- 351
 QY 167 GGTISAPGTAKNATTVGATENLRPSFSYADNINHYAQFSSRPTK-DGRKRDVWAPGT 225
 DB 352 PSSISGPGADSDHALTVGA-----VDSDDRAAYFTSAGPRGDNALKDLAAPGV 400
 QY 226 YILSARSSLPADSSFWANHDSKYAYNGTSMATPIVAGVAQUREHFVNRGRTPESSL 285
 DB 401 DIRAARQLAPGTYG-----YTSMSGTSMATPHVAGVAAALAEQHPMTGARLKDALM 453
 QY 286 KAA--LIAGAADVGLGYPNGNGMG--RYTLDSLVNAYNESALSTSQKATYFTTATAG 342
 DB 454 STSEQLDASVYQLAGRVSVPDAVGARVATGSGADGFHRWPHDADRPVTKVITYSNSSD 513
 QY 343 PKIKISLWSADAPASTASVTLVNDLVLITAP---NGTRYVGNDFSAF 388
 DB 514 TVEELSLAVGADAGV---ATLA---DTALTVAHGTAATTVGDSKAP 557

Search completed: March 31, 2004, 16:08:56
 Job time : 35.55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 47.1304 Seconds

(without alignment)
2595.843 Million cell updates/sec

Title: US-09-985-689a-3

Sequence: 1 NDVARGIVKADVAQNNGY.....EVQATNPSGPRSLAIYH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003s.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2263	100.0	433	5	AAW50082 Bacillus
2	2251	99.5	433	5	AAW50084 Bacillus
3	2239	98.9	433	2	AAW26274 Alkali-Pr
4	2239	98.9	433	2	AAW61495 Modified
5	2239	98.9	433	2	AAW95698 Bacillus
6	2239	98.9	433	3	AAW95698 Bacillus
7	2239	98.9	433	3	AAW95698 Bacillus
8	2239	98.9	433	3	AAW95698 Bacillus
9	2239	98.9	433	3	AAW95698 Bacillus
10	2239	98.9	433	3	AAW95698 Bacillus
11	2239	98.9	433	3	AAW95698 Bacillus
12	2239	98.9	433	3	AAW95698 Bacillus
13	2239	98.9	433	3	AAW95698 Bacillus
14	2239	98.9	433	3	AAW95698 Bacillus
15	2239	98.9	433	3	AAW95698 Bacillus
16	2239	98.9	433	3	AAW95698 Bacillus
17	2239	98.9	433	3	AAW95698 Bacillus
18	2239	98.9	433	3	AAW95698 Bacillus
19	2239	98.9	433	3	AAW95698 Bacillus
20	2239	98.9	433	3	AAW95698 Bacillus
21	2239	98.9	433	3	AAW95698 Bacillus
22	2239	98.9	433	3	AAW95698 Bacillus
23	2239	98.9	433	3	AAW95698 Bacillus
24	2239	98.9	433	3	AAW95698 Bacillus
25	2239	98.9	433	3	AAW95698 Bacillus

26	422.5	18.7	522	2	AAW24122	AAW24122 Pyrococcus
27	422.5	18.7	522	2	AAW94838	AAW94838 Hyperther
28	422.5	18.7	654	2	AAW24129	AAW24129 Pyrococcus
29	422.5	18.7	654	2	AAW94841	AAW94841 Hyperther
30	421.5	17.7	659	2	AAW24123	AAW24123 Protease
31	370.5	16.4	545	4	ABW99483	ABW99483 T. yonsei
32	346.5	15.3	1237	6	ABW11343	ABW11343 Protein e
33	345.5	15.3	1079	6	ABW1180	ABW1180 Transglut
34	345.5	15.2	520	2	ABW07391	ABW07391 Foreign P
35	343.5	15.2	520	2	AAW13666	AAW13666 Fragment
36	343.5	15.2	734	2	AAW13667	AAW13667 Streptomy
37	343.5	15.2	823	2	AAW13668	AAW13668 DHPA-mel
38	312	13.8	806	2	AAW27481	AAW27481 RP-III re
39	305	13.5	903	2	AAW87007	AAW87007 Hyperther
40	305	13.5	1398	2	AAW87008	AAW87008 Pyrococ
41	305	13.5	1398	2	AAW24124	AAW24124 Pyrococ
42	305	13.5	1398	2	AAW94839	AAW94839 W09856926
43	296.5	13.1	580	7	ADD24927	ADD24927 Xanthomon
44	295	13.0	519	6	ABP76735	ABP76735 Streptomy
45	295	13.0	19938	6	ABP76678	ABP76678 Streptomy

ALIGNMENTS

RESULT 1
ID AAW50082 standard; protein; 433 AA.

AAW50082;

12-AUG-2002 (first entry)

Bacillus sp D6-(FERM P1592) alkaline protease protein fragment.

Alkaline protease; detergent; laundry; bleaching; dishwasher.

Bacillus sp.

EP1209233-A2.

29-MAY-2002.

22-NOV-2001; 2001EP-00127851.

22-NOV-2000; 2000JP-00355166.

12-APR-2001; 2001JP-00114048.

(KROS) KAO CORP.

Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

Okuda M, Saeki K;

WPI; 2002-437518/47.

New modified alkaline proteases useful in detergent compositions.

Claim 5; Page 13-15; 25pp; English.

This invention describes novel Bacillus sp. alkaline proteases useful in

detergent compositions, especially in laundry, bleaching or automatic

dishwasher detergents. The novel proteases have an increased detergency *

(34 - 38) compared to prior art alkaline proteases (31 and 23%). This

sequence represents a fragment of the alkaline protease E-1 from Bacillus

sp strain D6-(FERM-P1592) described in the method of the invention

Query Match 100.0%; Score 2263; DB 5; Length 433;

Best Local Similarity 100.0%; Pred. No 2.4e-10; Gaps 0;

Matches 433; Conservative 0; Mismatches 0; Indels 0;

1 NDVARGIVKADVAQNNGYVAVDTGLDTGRDSSMHAFFKRTALVALGRIN 60

```

Db      1 NDVARGIVKADVQNNNGLYGGQVVAADTGLDTRGNDSSMHEARFGKITLALYLGRTN 60
Qy      61 NADDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGLPSNLNTLFSQAMNA 120
Db      61 NADDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGLPSNLNTLFSQAMNA 120
Qy      121 GARLHTNSWGAAPVNGATYANSRQVDEYVRNNDMTVLPFAAGNEGPNNGTISAPGTAKNAIT 180
Db      121 GARLHTNSWGAAPVNGATYANSRQVDEYVRNNDMTVLPFAAGNEGPNNGTISAPGTAKNAIT 180
Qy      181 VGATEMYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSLAAPDSFW 240
Db      181 VGATEMYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSLAAPDSFW 240
Qy      241 ANYSKYAYMGTSMTAPIVAGNVAQLREHFIKNGRITKPSLIKALTAGATDVLGYP 300
Db      241 ANYSKYAYMGTSMTAPIVAGNVAQLREHFIKNGRITKPSLIKALTAGATDVLGYP 300
Qy      301 SGDOGGRVTLDKSLNVAAYNEATALTGOKATYSFOTQAGKPLKISLWTDAPGSTAS 360
Db      301 SGDOGGRVTLDKSLNVAAYNEATALTGOKATYSFOTQAGKPLKISLWTDAPGSTAS 360
Qy      361 YTLVNDLDELVTLPNGOKTYGNDFSYPDNMDGRNNVENVFINAPOSSTYTIIEVOAYNV 420
Db      361 YTLVNDLDELVTLPNGOKTYGNDFSYPDNMDGRNNVENVFINAPOSSTYTIIEVOAYNV 420
Qy      421 PSGPQRFSLAIYH 433
Db      421 PSGPQRFSLAIYH 433

RESULT 2
AAMS0084
ID AAMS0084 standard; protein; 433 AA.
XX
XX AAMS0084;
XX
XX 12-AUG-2002 (first entry)
XX
DE Bacillus sp SD-521 (FERM BP-11162) alkaline protease protein fragment.
XX
XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX
XX Bacillus sp.
XX
XX EP1209233-A2.
XX
XX 29-MAY-2002.
XX
XX 22-NOV-2001; 2001EP-00127851.
XX
XX 22-NOV-2000; 2000JP-00355166.
XX
XX 12-APR-2001; 2001JP-00114048.
XX
XX (KAOS) KAO CORP.
XX
XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
XX
XX Okuda M, Saeki K;
XX
XX WPI; 2002-437518/47.
XX
XX New modified alkaline proteases useful in detergent compositions.
XX
XX Claim 5; Page 16-18; 25pp; English.
XX
XX This invention describes novel Bacillus sp. alkaline proteases useful in
XX
XX detergent compositions, especially in laundry, bleaching or automatic
XX
XX dishwasher detergents. The novel proteases have an increased detergency &
XX
XX (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX
XX sequence represents a fragment of the alkaline protease SD-521 from
XX
XX Bacillus sp strain SD-521 (FERM BP-11162) described in the method of the
XX
XX invention

```

```

XX      XX Sequence 433 AA;
SQ      Query Match      99.5%; Score 2251; DB 5; Length 433;
        Best Local Similarity 99.3%; Pred. No. 2.2e-169;
        Matches 430; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 NDVARGIVKADVQNNNGLYGGQVVAADTGLDTRGNDSSMHEARFGKITLALYLGRTN 60
Db      1 NDVARGIVKADVQNNNGLYGGQVVAADTGLDTRGNDSSMHEARFGKITLALYLGRTN 60
Qy      61 NADDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGLPSNLNTLFSQAMNA 120
Db      61 NADDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGLPSNLNTLFSQAMNA 120
Qy      121 GARLHTNSWGAAPVNGATYANSRQVDEYVRNNDMTVLPFAAGNEGPNNGTISAPGTAKNAIT 180
Db      121 GARLHTNSWGAAPVNGATYANSRQVDEYVRNNDMTVLPFAAGNEGPNNGTISAPGTAKNAIT 180
Qy      181 VGATEMYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSLAAPDSFW 240
Db      181 VGATEMYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSLAAPDSFW 240
Qy      241 ANYSKYAYMGTSMTAPIVAGNVAQLREHFIKNGRITKPSLIKALTAGATDVLGYP 300
Db      241 ANYSKYAYMGTSMTAPIVAGNVAQLREHFIKNGRITKPSLIKALTAGATDVLGYP 300
Qy      301 SGDOGGRVTLDKSLNVAAYNEATALTGOKATYSFOTQAGKPLKISLWTDAPGSTAS 360
Db      301 SGDOGGRVTLDKSLNVAAYNEATALTGOKATYSFOTQAGKPLKISLWTDAPGSTAS 360
Qy      361 YTLVNDLDELVTLPNGOKTYGNDFSYPDNMDGRNNVENVFINAPOSSTYTIIEVOAYNV 420
Db      361 YTLVNDLDELVTLPNGOKTYGNDFSYPDNMDGRNNVENVFINAPOSSTYTIIEVOAYNV 420
Qy      421 PSGPQRFSLAIYH 433
Db      421 PSGPQRFSLAIYH 433

RESULT 3
AAR26274
ID AAR26274 standard; protein; 433 AA.
XX
XX AAR26274;
XX
XX 05-FEB-1993 (first entry)
XX
XX Alkali resistance; surface active agent resistance; detergency improver.
XX
XX Alkali-protease Ya enzyme.
XX
XX Bacillus sp. Y.
XX
XX JP04197182-A.
XX
XX 16-JUL-1992.
XX
XX 28-NOV-1990; 90JP-00327110.
XX
XX 28-NOV-1990; 90JP-00327110.
XX
XX (LION) LION CORP.
XX
XX WPI; 1992-288440/35.
XX
XX N-PSDB; AAQ27516.
XX
XX DNA coding alkali-protease Ya enzyme - has good alkali and surfactant
XX
XX resistance and improves detergency.
XX
XX Claim 2; Page 1; 17pp; Japanese.
XX
XX The sequence is that of alkali-protease Ya enzyme which can be used in

```

CC the recombinant production of Ya enzyme. Ya enzyme is excellent in alkali
 CC resistance and surface active agent resistance and improves detergency
 XX
 SQ Sequence 433 AA;

Query Match 98.9%; Score 2239; DB 2; Length 433;
 Best Local Similarity 98.8%; Pred. No. 1.9e-168;
 Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGRNDSMEHAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGRNDSMEHAFRGKITALYALGRTN 60
 QY 61 NANDPNGHGTHTVAGSVTLGNALNKGNAPQANLVFQSIMDSSGGIGLPSNLNTLFSQAMNA 120
 DB 61 NANDPNGHGTHTVAGSVTLGNALNKGNAPQANLVFQSIMDSSGGIGLPSNLNTLFSQAMNA 120
 QY 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRITKPDVTAPGTITLSARSSLAPDSSFW 240
 DB 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRITKPDVTAPGTITLSARSSLAPDSSFW 240
 QY 241 ANYNSKYAYMGTSMAPIVAGNVAQLREHFIKNRGITPKPSLTKALIAAGTDVGLGYP 300
 DB 241 ANYNSKYAYMGTSMAPIVAGNVAQLREHFIKNRGITPKPSLTKALIAAGTDVGLGYP 300
 QY 301 SGOQGWGRVTLDSLVAVAYNEATALTGQKATYSFOAQGRPKLISLWTDAPGSTTAS 360
 DB 301 SGOQGWGRVTLDSLVAVAYNEATALTGQKATYSFOAQGRPKLISLWTDAPGSTTAS 360
 QY 361 YTLVNDLIDLIVITAPNGQKYVGNDFSPYDNNMDGRNNVENFINAPQSGTITIEVOAYNV 420
 DB 361 YTLVNDLIDLIVITAPNGQKYVGNDFSPYDNNMDGRNNVENFINAPQSGTITIEVOAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433

RESULT 4
 AAM61495 standard; protein; 433 AA.
 XX
 AC AAM61495;
 XX
 DT 06-NOV-1998 (first entry)
 XX
 DE Modified Bacillus lichen protease.
 XX
 KM Bacillus lichen Y protease; polyethylene glycol; PEG; soap;
 KM methoxypolyethyleneglycol; MPEG; skin; hair care product; cosmetic;
 KM lipstick; hair gel; sun oil; shampoo; hair dye; insect repellent.
 OS Bacillus sp.
 XX
 FH Key
 FT Modified-site 1..433 Location/Qualifiers
 FT /note="The enzyme is modified by methoxypolyethyleneglycol
 FT molecules covalently attached to the N-terminal amino
 FT group and to fourteen unspecified amino groups of lysine
 FT residues present on the surface of the enzyme"
 XX
 EN MO9830682-A1.
 XX
 PD 16-UTL-1998.
 XX
 PF 12-JAN-1998; 98WO-DK000015.
 XX
 PR 10-JAN-1997; 97DK-00000038.
 PR 25-JUN-1997; 97DK-00000754.

XX (NOVO) NOVO-NORDISK AS.
 PA
 XX
 PI Olsen AA, Prento A;
 XX
 DR WPI; 1998-399132/34.
 XX
 PT New enzyme modified by attachment of many polymeric molecules - useful in
 PT skin and hair care products, has reduced tendency to cause sensitisation
 PT and increased stability.
 PS Claim 16; Page 44-45; 56pp; English.
 XX

CC The present sequence represents the Bacillus lichen Y protease. The
 CC invention claims for enzymes covalently modified on their surface by the
 CC attachment of a large number of small polymeric molecules, e.g.
 CC polyethylene glycols (PEG). The polymeric molecules are coupled to the N-
 CC terminal amino group and the amino groups of lysine residues found on the
 CC surface of the enzyme. In the example given, the lichen Y protease was
 CC modified using methoxypolyethyleneglycol (MPEG) as the polymeric molecule.
 CC The N-terminal amino group and the amino groups of the fourteen lysine
 CC residues present on the surface of the lichen Y protease were modified.
 CC Modification of the enzymes increases the stability and/or reduces the
 CC sensitising potential (allergenicity) of the enzyme, without
 CC significantly reducing enzymatic activity. Also, using a large number of
 CC relatively small polymeric molecules, rather than a few very large ones,
 CC provides a more even effect with reduced activity loss. The modified
 CC enzymes are claimed to be useful as components of a wide range of skin
 CC and hair care products, e.g. soaps, cosmetics, creams, gels, lipsticks,
 CC hair gels, sun oils, shampoos, hair dyes, insect repellants, etc
 XX

Query Match 98.9%; Score 2239; DB 2; Length 433;
 Best Local Similarity 98.8%; Pred. No. 1.9e-168;
 Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGRNDSMEHAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGRNDSMEHAFRGKITALYALGRTN 60
 QY 61 NANDPNGHGTHTVAGSVTLGNALNKGNAPQANLVFQSIMDSSGGIGLPSNLNTLFSQAMNA 120
 DB 61 NANDPNGHGTHTVAGSVTLGNALNKGNAPQANLVFQSIMDSSGGIGLPSNLNTLFSQAMNA 120
 QY 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAAPVNGAYTANSRQVDEYVANNMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRITKPDVTAPGTITLSARSSLAPDSSFW 240
 DB 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGRITKPDVTAPGTITLSARSSLAPDSSFW 240
 QY 241 ANYNSKYAYMGTSMAPIVAGNVAQLREHFIKNRGITPKPSLTKALIAAGTDVGLGYP 300
 DB 241 ANYNSKYAYMGTSMAPIVAGNVAQLREHFIKNRGITPKPSLTKALIAAGTDVGLGYP 300
 QY 301 SGOQGWGRVTLDSLVAVAYNEATALTGQKATYSFOAQGRPKLISLWTDAPGSTTAS 360
 DB 301 SGOQGWGRVTLDSLVAVAYNEATALTGQKATYSFOAQGRPKLISLWTDAPGSTTAS 360
 QY 361 YTLVNDLIDLIVITAPNGQKYVGNDFSPYDNNMDGRNNVENFINAPQSGTITIEVOAYNV 420
 DB 361 YTLVNDLIDLIVITAPNGQKYVGNDFSPYDNNMDGRNNVENFINAPQSGTITIEVOAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433

RESULT 5
 AAM95698 standard; protein; 433 AA.
 ID AAM95698

XX	AAW95698;
AC	
XX	16-JUN-1999 (first entry)
DT	
XX	Bacillus sp. Lion Y protease.
DE	
XX	PD498; subtilisin; Lion Y; protease; skin-care; feed; additive; soap;
KW	cosmetic; hair dye; sunscreens; acne; antiperspirants; insect repellent;
KM	deodorant; detergent; food; breadmaking; textile-treating; oral; dermal;
KW	pharmaceutical; agrochemical.
XX	
OS	Bacillus sp.
XX	
PN	WO900485-A1.
XX	
PD	07-JAN-1999.
XX	
PF	22-JUN-1998; 98WO-DK000270.
XX	
PR	25-JUN-1997; 97DK-00000753.
PR	07-JUL-1997; 97US-0051830P.
XX	
PA	(NOVO) NOVO-NORDISK AS.
XX	
PI	Olsen AA, Fatum TM, Deussen H, Roggen EL;
XX	
DR	WPI; 1999-095735/08.
XX	
PT	New modified polypeptide with attached low molecular weight polymer - has
PT	reduced respiratory allergenicity, useful in skin care products,
PT	detergents, as food additives or textile-treating compositions.
XX	
PS	Claim 10; Page 48-49; 60pp; English.
XX	
CC	The sequence is that of Lion Y protease. This can be used as an active
CC	ingredient: (i) in personal care products (especially skin-care products
CC	such as soaps, cosmetics, hair dyes, sunscreens, anti-acne products,
CC	antiperspirants, insect repellants or deodorants); (ii) in detergents (as
CC	laundry, dishwashing or hard-surface cleaners); (iii) food or feed
CC	additives (e.g. for breadmaking); (iv) in textile-treating compositions,
CC	or (v) in oral or dermal pharmaceuticals and agrochemicals
XX	
XX	Sequence 433 AA;

QY	361	YTLVNDLDTLTANNGKRYGVNDSPYDNNMDGRNNVENVFIAPOSQTYTLEVPAYNV	420
DB	361	YTLVNDLDTLTANNGKRYGVNDSPYDNNMDGRNNVENVFIAPOSQTYTLEVPAYNV	420
QY	421	PSGPFQSFSLATVH	433
DB	421	PSGPFQSFSLATVH	433
RESULT 6			
ID	AAV69207	standard; protein; 433 AA.	
XX	AAV69207;		
DT	30-MAY-2000	(first entry)	
XX			
DE	Amino acid sequence of protease Lion Y.		
XX			
KW	Protease Lion Y; polypeptide-polymer conjugate; washing performance;		
KM	respiratory allergenicity; allergic reaction; detergent formulation;		
KW	laundry; dishwashing; hard surface cleaner; agricultural chemical;		
KM	skin care; cosmetic; oral pharmaceutical; dental pharmaceutical;		
XX	textile processing.		
OS	Bacillus sp.		
PN	WO200004138-A1.		
XX			
PD	27-JAN-2000.		
XX			
PF	16-JUL-1999; 99WO-DK000406.		
XX			
PR	17-JUL-1998; 98DK-00000951.		
PA	(NOVO) NOVO-NORDISK AS.		
XX			
PI	Bauditz P, Fatum TM, Olsen AA, Deussen H, Petersen DA;		
XX			
DR	WPI; 2000-195024/17.		
PT	New polypeptide-polymer conjugate, particularly enzyme conjugate, useful		
XX	in detergent formulations.		
PS	Disclosure: Page 74-76; 79pp; English.		
XX			
CC	The present sequence represents a protease Lion Y protein. The protein		
CC	may be used to produce the conjugates of the invention. The specification		
CC	describes polypeptide-polymer conjugates which have improved washing		
CC	performance and reduced respiratory allergenicity compared with the		
CC	unconjugated polypeptide. The polymer provides, in water, a conformation		
CC	that shields the molecular surface effectively, preventing association of		
CC	antibodies that can induce an allergic reaction. The conjugates are used		
CC	in industrial compositions, particularly detergent formulations (laundry,		
CC	dishwashing or hard surface cleaners), but also in agricultural		
CC	chemicals, skin care products (cosmetics and toiletries), oral and dental		
CC	pharmaceuticals, or textile processing and treatment compositions		
XX			
SQ	Sequence 433 AA;		
QY	Query Match	98.9%; Score 2239; DB 3; Length 433;	
	Best Local Similarity	98.8%; Pred. No. 1.9e-168;	
	Matches 428; Conservative	2; Mismatches 3; Indels 0; Gaps 0;	
DB	1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMHEAFQKITATLYALGRTN	60	
QY	61 NANDNGHGTVAAGSYGNALNKMAQOANLVFQSIDSSGGGLGLSNLNTLPSQAMNA	120	
DB	61 NASDPNGHGTVAAGSYGNALNKMAQOANLVFQSIDSSGGGLGLSNLNTLPSQAMNA	120	
QY	121 GAEHTHSWGPVNGAVTANSRQVDEYVRNNDMTVLEAAGNEGPNSGTISAPGTAKAAIT	180	

Db 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGR:KPDVTAGTFLISARSSLA PDSSFW 240
 Db 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGR:KPDVTAGTFLISARSSLA PDSSFW 240
 QY 241 ANYNSKYAYMGSTSMATPIVAGNVAQLEHFIKNGITPKPSLIIKALIAAGATDVGLGYP 300
 Db 241 ANYNSKYAYMGSTSMATPIVAGNVAQLEHFIKNGITPKPSLIIKALIAAGATDVGLGYP 300
 QY 301 SDDQGWGRVTLDKSLNVAAYVNEATLALTTGOKATYSFQOAGKPKISLWTDAPGSTTAS 360
 Db 301 SDDQGWGRVTLDKSLNVAAYVNEATLALTTGOKATYSFQOAGKPKISLWTDAPGSTTAS 360
 QY 361 YTLVNDLDELVTAPNGQKTVGNDFSYPYDNNMDGRNNEVFINAPQSGTYTIEVOAYNV 420
 Db 361 YTLVNDLDELVTAPNGQKTVGNDFSYPYDNNMDGRNNEVFINAPQSGTYTIEVOAYNV 420
 QY 421 PSGPQRFSLAIYH 433
 Db 421 PSGPQRFSLAIYH 433
 RESULT 7
 ID AAY44619 standard, protein, 433 AA.
 AC AAY44619;
 DT 07-APR-2000 (first entry)
 DE Bacillus lion y enzyme.
 XX Lion Y enzyme; protease; allergic response; industrial composition;
 XX co-polymer; ethylene oxide; EO; propylene oxide; PO; conjugate;
 XX allergenicity; detergent; cosmetic; toiletries; textile treatment;
 XX agrochemical; pharmaceutical; food; feed additive.
 OS Bacillus sp.
 XX WO9967370-A1.
 XX 29-DEC-1999.
 XX 23-JUN-1999; 99WO-DK000359.
 XX 23-JUN-1998; 98DK-00000809.
 XX (NOVO) NOVO-NORDISK AS.
 XX Deussen H, Olsen AA, Fatum TM, Roggen EL;
 XX WPI; 2000-136981/12.
 XX New conjugate of polypeptide, especially an enzyme, with copolymer of
 XX ethylene oxide and propylene oxide, used in e.g. cleaning compositions,
 XX has reduced allergenicity.
 PS Claim 10; Page 56-57; 62pp; English.
 XX The present sequence is a Bacillus lion y enzyme, which is a protease
 XX capable of inducing an allergic response upon inhalation. The enzyme can
 XX be covalently coupled to a co-polymer comprising ethylene oxide (EO) and
 XX propylene oxide (PO) to reduce its allergenicity. This enzyme-polymer
 XX conjugate can be used in industrial compositions such as detergents,
 XX cosmetics, toiletries, textile treatment compositions, agrochemicals,
 XX oral and dermal pharmaceuticals and food and feed additives
 XX Sequence 433 AA;
 Query Match 98.9%; Score 2239; DB 3; Length 433;
 Best Local Similarity 98.8%; Pred. No. 1.9e-168;

Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADYAKONNNGLYGQGVAVAVADTGLDTRGNDSSWHEAFRGKITLVALGRTN 60
 Db 1 NDVARGIVKADYAKONNNGLYGQGVAVAVADTGLDTRGNDSSWHEAFRGKITLVALGRTN 60
 QY 61 NADPNHGHTHVAAGSVLGNALNKGMAPOANLVFQSIMDSGGLGGLPSNLNTLFGQAMNA 120
 Db 61 NADPNHGHTHVAAGSVLGNALNKGMAPOANLVFQSIMDSGGLGGLPSNLNTLFGQAMNA 120
 QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGPNSGTISAPGTAKNAIT 180
 Db 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGR:KPDVTAGTFLISARSSLA PDSSFW 240
 Db 181 VGATENYRPSFGSIADNPNHIAQFSSRGATRDGR:KPDVTAGTFLISARSSLA PDSSFW 240
 QY 241 ANYNSKYAYMGSTSMATPIVAGNVAQLEHFIKNGITPKPSLIIKALIAAGATDVGLGYP 300
 Db 241 ANYNSKYAYMGSTSMATPIVAGNVAQLEHFIKNGITPKPSLIIKALIAAGATDVGLGYP 300
 QY 301 SDDQGWGRVTLDKSLNVAAYVNEATLALTTGOKATYSFQOAGKPKISLWTDAPGSTTAS 360
 Db 301 SDDQGWGRVTLDKSLNVAAYVNEATLALTTGOKATYSFQOAGKPKISLWTDAPGSTTAS 360
 QY 361 YTLVNDLDELVTAPNGQKTVGNDFSYPYDNNMDGRNNEVFINAPQSGTYTIEVOAYNV 420
 Db 361 YTLVNDLDELVTAPNGQKTVGNDFSYPYDNNMDGRNNEVFINAPQSGTYTIEVOAYNV 420
 QY 421 PSGPQRFSLAIYH 433
 Db 421 PSGPQRFSLAIYH 433
 RESULT 8
 ID AAM50083 standard, protein, 433 AA.
 AC AAM50083;
 DT 12-AUG-2002 (first entry)
 DE Bacillus sp Y-(FERM BP-1029) alkaline protease protein fragment.
 XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
 OS Bacillus sp.
 XX EP1209233-A2.
 XX 29-MAY-2002.
 XX 22-NOV-2001; 2001EP-00127851.
 XX 22-NOV-2000; 2000JP-00355166.
 XX 12-APR-2001; 2001JP-00114048.
 XX (KAO) KAO CORP.
 XX Harada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
 XX Okuda M, Saeki K;
 XX WPI; 2002-437518/47.
 XX New modified alkaline proteases useful in detergent compositions.
 XX Claim 5; Page 15-16; 25pp; English.
 XX This invention describes novel Bacillus sp. alkaline proteases useful in
 XX detergent compositions, especially in laundry, bleaching or automatic
 XX dishwasher detergents. The novel proteases have an increased detergent
 XX (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This

CC sequence represents a fragment of the alkaline protease Ya from *Bacillus*
CC sp strain Y-(FERM BP-1029) described in the method of the invention
XX
SQ Sequence 433 AA:
Query Match 98.7%; Score 2234; DB 5; Length 433;
Best Local Similarity 98.6%; Pred. No. 4.8e-168;
Matches 427; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNNVGLYGQGVAVADTGLDGRNDSMHEAFRGKITLVALGRIN 60
DB 1 NDVARGIVKADVAQNNVGLYGQGVAVADTGLDGRNDSMHEAFRGKITLVALGRIN 60
QY 61 NADPENGHTHVASVILGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAMNA 120
DB 61 NADPENGHTHVASVILGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAMNA 120
QY 121 GARHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
DB 121 GARHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKIPDYTAFTLLSARSSLAPSSFW 240
DB 181 VGATENYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKIPDYTAFTLLSARSSLAPSSFW 240
QY 241 ANYSKYAVMGTSMAPIVAGNVAQLREHFIKNRGITPKPSLIKALLAGATDVLGYP 300
DB 241 ANYSKYAVMGTSMAPIVAGNVAQLREHFIKNRGITPKPSLIKALLAGATDVLGYP 300
QY 301 SGGDQGWGRVTLDSLVNAVYNNETALTGGOKATYSFQTAQKPLKISLWTDAPGSTTAS 360
DB 301 SGGDQGWGRVTLDSLVNAVYNNETALTGGOKATYSFQTAQKPLKISLWTDAPGSTTAS 360
QY 361 YTLVNDLDLVITAPNGQKTVGNDFSYPYDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
DB 361 YTLVNDLDLVITAPNGQKTVGNDFSYPYDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
QY 421 PSGBPQFSLAIIVH 433
DB 421 PSGBPQFSLAIIVH 433
RESULT 9
AA89548
ID AA89548 standard; protein; 636 AA.
AC AA89548;
XX
DT 12-APR-1999 (first entry)
XX
DE *Bacillus* sp. alkaline protease Y.
XX
KM Alkaline protease Y, detergent; surfactant; leather processing;
XX debittering; flavour.
XX
OS *Bacillus* sp.
XX
PN WO9856927-A2.
XX
PD 17-DEC-1998.
XX
PF 09-JUN-1998; 98WO-US012005.
XX
PR 12-JUN-1997; 97US-00873479.
XX
PA (NOVO) NOVO NORDISK BIOTECH INC.
XX
PI Sioma A, Christianson L;
XX
DR WPI, 1999-080908/07.
XX
PT Novel protease from *Bacillus subtilis* LC20 - useful in laundry and
XX dishwashing detergents and for leather processing.

XX
XX Claim 3; Page 55-56; 77pp; English.
CC This is the amino acid sequence of a *Bacillus* sp. alkaline protease Y
CC that is said to have good alkali and surfactant resistance and improved
CC detergency. It shows 77% identity to a newly isolated protease (see
CC AA89547) of *Bacillus* sp. JP170 (NCIB 12513). The invention provides
CC vectors, recombinant host cells and methods for the recombinant
CC production of such proteases. The protease are used in laundry and
CC dishwashing detergents, for institutional and industrial cleaning, and
CC for leather processing, as well as for debittering and enhancing the
CC degree of hydrolysis of protein hydrolysates, for flavour development
CC through hydrolysis of peptides, degradation of undesired peptides and in
CC enzymatic synthesis of peptides. They have enhanced stability towards
CC oxidation under alkaline conditions, e.g. towards bleaching agents of the
CC peroxyl type. The invention also provides mutant cells in which the
CC protease activity is diminished. Such cells can be used for the
CC production of heterologous recombinant proteins
XX
SQ Sequence 636 AA:
Query Match 98.5%; Score 2230; DB 2; Length 636;
Best Local Similarity 98.6%; Pred. No. 1.7e-167;
Matches 427; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNNVGLYGQGVAVADTGLDGRNDSMHEAFRGKITLVALGRIN 60
DB 204 NDVARGIVKADVAQNNVGLYGQGVAVADTGLDGRNDSMHEAFRGKITLVALGRIN 263
QY 61 NADPENGHTHVASVILGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAMNA 120
DB 61 NADPENGHTHVASVILGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAMNA 120
QY 121 GARHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
DB 121 GARHTNSWGAAPVNGAYTANSRQVDEYVNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKIPDYTAFTLLSARSSLAPSSFW 240
DB 181 VGATENYRPSFGSIADNPNIHIAQFSSRGATRDGRIRKIPDYTAFTLLSARSSLAPSSFW 240
QY 241 ANYSKYAVMGTSMAPIVAGNVAQLREHFIKNRGITPKPSLIKALLAGATDVLGYP 300
DB 241 ANYSKYAVMGTSMAPIVAGNVAQLREHFIKNRGITPKPSLIKALLAGATDVLGYP 300
QY 301 SGGDQGWGRVTLDSLVNAVYNNETALTGGOKATYSFQTAQKPLKISLWTDAPGSTTAS 360
DB 301 SGGDQGWGRVTLDSLVNAVYNNETALTGGOKATYSFQTAQKPLKISLWTDAPGSTTAS 360
QY 361 YTLVNDLDLVITAPNGQKTVGNDFSYPYDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
DB 361 YTLVNDLDLVITAPNGQKTVGNDFSYPYDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
QY 421 PSGBPQFSLAIIVH 433
DB 421 PSGBPQFSLAIIVH 433
RESULT 10
AA850086
ID AA850086 standard; protein; 433 AA.
AC AA850086;
XX
DT 12-AUG-2002 (first entry)
XX
DE *Bacillus* sp alkaline protease protein A-2 fragment.
XX
KM Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX
OS *Bacillus* sp.
XX
PN EPI209233-A2.

```

XX PD 29-MAY-2002.
XX PF 22-NOV-2001; 2001EP-00127851.
XX PR 22-NOV-2000; 2000UP-00355166.
XX PR 12-APR-2001; 2001JP-00114048.
XX PA (KAOS ) KAO CORP.
XX PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
XX PI Okuda M, Saeki K;
XX DR MPI; 2002-437518/47.
XX FT New modified alkaline proteases useful in detergent compositions.
XX PS Claim 5; Page 20-21; 25pp; English.
XX CC This invention describes novel Bacillus sp. alkaline proteases useful in
XX CC detergent compositions, especially in laundry, bleaching or automatic
XX CC dishwasher detergents. The novel proteases have an increased detergency %
XX CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX CC sequence represents a fragment of the alkaline protease A-2 from Bacillus
XX CC sp NC1B12513 described in the method of the invention
XX SQ Sequence 433 AA;
XX
Query Match 90.1%; Score 2040; DB 5; Length 433;
Best Local Similarity 88.9%; Pred. No. 1e-152;
Matches 385; Conservative 24; Mismatches 24; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNNNGLYGGQGVVAADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
DB 1 NDVARGIVKADVAQNNNGLYGGQGVVAADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
QY 61 NANDPNGHGHVAVSVGNALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLFSGQAWNA 120
DB 61 NANDPNGHGHVAVSVGNALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLFSGQAWNA 120
QY 121 GARIHNSWGAPEVNGATYANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
DB 121 GARIHNSWGAPEVNGATYANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
QY 121 GARIHNSWGAPEVNGATYANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
DB 121 GARIHNSWGAPEVNGATYANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
QY 181 VGATENLRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 240
DB 181 VGATENLRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 240
QY 241 ANYSKXAYVGTSMATPIVAGNVAQLREHIFNRGITKPSLIKALITAGATDVGLYP 300
DB 241 ANYSKXAYVGTSMATPIVAGNVAQLREHIFNRGITKPSLIKALITAGATDVGLYP 300
QY 241 ANHDSKXAYVGTSMATPIVAGNVAQLREHIFNRGITKPSLIKALITAGATDVGLYP 300
DB 241 ANHDSKXAYVGTSMATPIVAGNVAQLREHIFNRGITKPSLIKALITAGATDVGLYP 300
QY 301 SSGQSGRVTLDLSLNAVAYNEATALTGGKATYSQTQAGKLSLVTWTPAPSGTTAS 360
DB 301 SSGQSGRVTLDLSLNAVAYNEATALTGGKATYSQTQAGKLSLVTWTPAPSGTTAS 360
QY 361 YTLVNDLDTLVTAPNGQKYGNDFSYEDNNMDGRNNVENVFINAQSGTITVEQAYNV 420
DB 361 YTLVNDLDTLVTAPNGQKYGNDFSYEDNNMDGRNNVENVFINAQSGTITVEQAYNV 420
QY 421 PSGPQRFSLAIVH 433
DB 421 PSGPQRFSLAIVH 433

```

RESULT 11

```

AA89547
ID AA89547 standard; protein; 641 AA.
XX AC AA89547;
XX DT 12-APR-1999 (first entry)
XX

```

```

DE Bacillus JP170 protease.
XX KM Protease, detergent; surfactant; leather processing; debittering;
XX KM flavour.
XX OS Bacillus sp.
XX FH Key
XX FT Peptide 1..33
XX FT Region /note="signal peptide"
XX FT Protein /note="prepro region"
XX FT Protein /note="mature protein"
XX PM MO9856927-A2.
XX PD 17-DEC-1998.
XX PF 09-JUN-1998; 98WO-US012005.
XX PR 12-JUN-1997; 97US-00873479.
XX PA (NOVO ) NOVO NORDISK BIOTECH INC.
XX PI Sloma A, Christianson L;
XX PI MPI; 1999-080908/07.
XX DR N-PSDB; AA892382.
XX FT Novel protease from Bacillus subtilis LC20 - useful in laundry and
XX FT dishwashing detergents and for leather processing.
XX PS Claim 7; Page 53-54; 77pp; English.
XX
This is the amino acid sequence of a novel protease of Bacillus sp. JP170
(CC NC1B 12513), as deduced from the nucleotide sequence of an isolated gene
(CC (see AA892382). The entire protein, including the signal peptide and
(CC prepro region, has 77% identity to alkaline protease Y (see AA89548)
(CC from Bacillus. The invention provides vectors, recombinant host cells and
(CC methods for the recombinant production of the protease. The protease is
(CC used in laundry and dishwashing detergents, for institutional and
(CC industrial cleaning, and for leather processing, as well as for
(CC debittering and enhancing the degree of hydrolysis of protein
(CC hydrolyses, for flavour development through hydrolysis of proteins,
(CC degradation of undesired peptides and in enzymatic synthesis of peptides.
(CC It has enhanced stability towards oxidation under alkaline conditions,
(CC e.g. towards bleaching agents of the peroxy type. The invention also
(CC provides mutant cells in which the protease activity is diminished. Such
(CC cells can be used for the production of heterologous recombinant proteins
XX
SQ Sequence 641 AA;
XX
Query Match 90.1%; Score 2040; DB 2; Length 641;
Best Local Similarity 88.9%; Pred. No. 1.8e-152;
Matches 385; Conservative 24; Mismatches 24; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNNNGLYGGQGVVAADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
DB 1 NDVARGIVKADVAQNNNGLYGGQGVVAADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
QY 209 NDVARGIVKADVAQNNNGLYGGQGVVAADTGLDTGRNDSMHEAFRGKITALYALGRTN 268
DB 209 NDVARGIVKADVAQNNNGLYGGQGVVAADTGLDTGRNDSMHEAFRGKITALYALGRTN 268
QY 61 NANDPNGHGHVAVSVGNALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLFSGQAWNA 120
DB 61 NANDPNGHGHVAVSVGNALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLFSGQAWNA 120
QY 269 NANDPNGHGHVAVSVGNALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLFSGQAWNA 328
DB 269 NANDPNGHGHVAVSVGNALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLFSGQAWNA 328
QY 121 GARIHNSWGAPEVNGATYANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
DB 121 GARIHNSWGAPEVNGATYANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
QY 329 GARIHNSWGAPEVNGATYANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 388
DB 329 GARIHNSWGAPEVNGATYANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 388
QY 181 VGATENLRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 240
DB 181 VGATENLRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 240
QY 389 VGATENLRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 448
DB 389 VGATENLRPSFGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFTLSARSSILAPSSFW 448

```

```

QY 241 ANNSKAYVGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKALIAGATDVLGY 300
DB 449 ANHDSKAYVGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKALIAGADVLGFP 508
QY 301 SGDOGWRVTLDKSLNVAAYVNEATALTGOKATYSFOTOGKPKLSLWTDAPGSTAS 360
DB 509 NGNGQWGRVTLDKSLNVAAYVNEATALTGOKATYSFOTOGKPKLSLWSDAPGSTAS 568
QY 361 YTLVNDLDLVITAPNGOKTVGNDPSYFDNNMDGRNNVENVFIMAPQSGTYTIEVOAYN 420
DB 569 LTLVNDLDLVITAPNGOKTVGNDPSYFDNNMDGRNNVENVFIMAPQSGTYTIEVOAYN 628
QY 421 PSQGPQFSLAIYH 433
DB 629 FVSPQFSLAIYH 641

RESULT 12
ID AAM50085 standard; protein; 434 AA.
AC AAM50085;
XX 12-AUG-2002 (first entry)
DT 12-AUG-2002 (first entry)
DE Bacillus sp alkaline protease protein A-1 fragment.
XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX Bacillus sp.
XX EPI209233-A2.
XX 29-MAY-2002.
XX 22-NOV-2001; 2001EP-00127851.
XX 22-NOV-2000; 2000JP-00355166.
XX 12-APR-2001; 2001JP-00114048.
XX (KAOS ) KAO CORP.
XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
PI Okuda M, Saeki K;
DR WPI; 2002-437518/47.
XX New modified alkaline proteases useful in detergent compositions.
XX Claim 5; Page 18-19; 25pp; English.
XX This invention describes novel Bacillus sp. alkaline proteases useful in
XX detergent compositions, especially in laundry, bleaching or automatic
XX dishwasher detergents. The novel proteases have an increased detergency %
XX (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX sequence represents a fragment of the alkaline protease A-1 from Bacillus
XX sp NCIB1289 described in the method of the invention
XX
XX Sequence 434 AA:
Query Match 89.3%; Score 2020.5; DB 5; Length 434;
Best Local Similarity 88.2%; Pred. No. 3.7e-151;
Matches 383; Conservative 28; Mismatches 22; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVQNNYGLYGQGVVAVADTGLDTRGNSMHEARRGKITLALYALGRN 60
DB 1 NDVARGIVKADVQNNYGLYGQGVVAVADTGLDTRGNSMHEARRGKITLALYALGRN 60
QY 61 NADPNHGTHVAGSVLGNAL-NKGNAPQANLVFQSIIMDSGGIGLPSNLNTLFSQAWN 119
DB 61 NADPNHGTHVAGSVLGNAL-NKGNAPQANLVFQSIIMDSGGIGLPSNLNTLFSQAWN 120
QY 120 AGARIHNSMGAPVNGAYTANSQVDEYVANNMTYLFAGNGRPNSGTISAPGTAKNAI 179

```

```

DB 121 AGARIHNSMGAPVNGAYTANSQVDEYVANNMTYLFAGNGRPNSGTISAPGTAKNAI 180
QY 180 TVGATENVRPFGSIAINPNHIAQFSSRGATPDGRIRKPDYTAGQFTLSARSSIAAPSSF 239
DB 181 TVGATENVRPFGSIAINPNHIAQFSSRGATPDGRIRKPDYTAGQFTLSARSSIAAPSSF 240
QY 240 WANNYSKAYVGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKALIAGATDVLGY 299
DB 241 WANNYSKAYVGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKALIAGATDVLGY 300
QY 300 PSQGPQFSLAIYH 433
DB 301 PSQGPQFSLAIYH 433
QY 360 SYTLVNDLDLVITAPNGOKTVGNDPSYFDNNMDGRNNVENVFIMAPQSGTYTIEVOAYN 419
DB 361 SYTLVNDLDLVITAPNGOKTVGNDPSYFDNNMDGRNNVENVFIMAPQSGTYTIEVOAYN 420
QY 420 VPSGPQFSLAIYH 433
DB 421 VPSGPQFSLAIYH 434

RESULT 13
ID AAM50081 standard; protein; 434 AA.
AC AAM50081;
XX 12-AUG-2002 (first entry)
DT 12-AUG-2002 (first entry)
DE Bacillus sp KSM-KP9860 alkaline protease protein fragment.
XX Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX Bacillus sp.
XX EPI209233-A2.
XX 29-MAY-2002.
XX 22-NOV-2001; 2001EP-00127851.
XX 22-NOV-2000; 2000JP-00355166.
XX 12-APR-2001; 2001JP-00114048.
XX (KAOS ) KAO CORP.
XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
PI Okuda M, Saeki K;
DR WPI; 2002-437518/47.
XX New modified alkaline proteases useful in detergent compositions.
XX Claim 5; Page 12-13; 25pp; English.
XX This invention describes novel Bacillus sp. alkaline proteases useful in
XX detergent compositions, especially in laundry, bleaching or automatic
XX dishwasher detergents. The novel proteases have an increased detergency %
XX (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX sequence represents a fragment of the alkaline protease KP9860 from
XX Bacillus sp strain KSM-KP9860 described in the method of the invention
XX
XX Sequence 434 AA:
Query Match 88.7%; Score 2006.5; DB 5; Length 434;
Best Local Similarity 88.2%; Pred. No. 4.7e-150;
Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVQNNYGLYGQGVVAVADTGLDTRGNSMHEARRGKITLALYALGRN 60

```


XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX WPI: 1999-287736/27.
DR N-PSDB; AAX37279.
XX

PT Alkali protease from *Bacillus* used in washing powders.

PS Disclosure; Page 63-68; 71pp; Japanese.

XX The invention relates to alkaline proteases produced by strains of
CC *Bacillus*. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidising agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)
XX

SQ Sequence 640 AA;

Query Match 88.2%; Score 1995.5; DB 2; Length 640;
Best Local Similarity 87.6%; Pred. No. 6e-149;

Matches 380; Conservative 26; Mismatches 25; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLVGGQVAVADTGLDTGRNDSMHEAFKGTALYALGRIN 60
DB NDVARGIVKADVAQSSYGLVGGQIVAVADTGLDTGRNDSMHEAFKGTALYALGRIN 266
QY 61 NADPNHGHGVASVIGN-ALNKNMKAPOANLVFQSIMDSGGIGLPSNYLTLPQAWN 119
DB NADPNHGHGVASVIGNSGNSTKNMKAPOANLVFQSIMDSGGIGLPSNYLTLPQAWN 326
QY 120 AGAIIHNSWGAIPNGAYITANSROYDEYRNNMTVTFAAGTEGNSGTTISAPGTAKNAI 179
DB AGAIIHNSWGAIPNGAYITANSROYDEYRNNMTVTFAAGTEGNSGTTISAPGTAKNAI 327
QY 180 TVGATENYRPSFGSIADNPNIHIAQFSSRGATRDRIKPDVTAPGTFLSARSSLAPDSF 239
DB TVGATENYRPSFGSIADNPNIHIAQFSSRGATRDRIKPDVTAPGTFLSARSSLAPDSF 387
QY 240 WANYNSKAYMGTSNATPIVAGNVAQLREHFIKNGITPKPSLIKALILAGATDVGLGY 299
DB WANYNSKAYMGTSNATPIVAGNVAQLREHFIKNGITPKPSLIKALILAGATDVGLGY 447
QY 300 PSGDQGWGRVTLDSLVNAVYNEATALTGQKATYSFQTQAGKPLKISLVMTDAPGSTTA 359
DB PSGDQGWGRVTLDSLVNAVYNEATALTGQKATYSFQTQAGKPLKISLVMTDAPGSTTA 507
QY 360 SYTLVNDLDLVITAPNGQKTVGNDSPYDNNMDGRNNVENVFINAPQSGTYTIEVOAYN 419
DB SYTLVNDLDLVITAPNGQKTVGNDSPYDNNMDGRNNVENVFINAPQSGTYTIEVOAYN 567
QY 420 VPSGPQRFSLAIVH 433
DB VPSGPQRFSLAIVH 627
VPIVGPQNFSLAIVH 640

Search completed: March 31, 2004, 16:04:30
Job time: 48.1304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:59:39 ; Search time 13.9892 Seconds

(without alignments)
1597.947 Million cell updates/sec

Title: US-09-985-689a-3

Perfect score: 2263

Sequence: 1 NDVARGIVKADVAQNNYGLV.....EYQAVNPSPGPRFSLATVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: Issued Parents AA:*

```
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCITUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2239	98.9	433	4	US-09-104-623A-4
2	2239	98.9	433	4	US-09-019-532-4
3	2239	98.9	433	4	US-09-338-746-4
4	2239	98.9	433	2	US-08-873-479-43
5	2040	90.1	641	2	US-08-873-479-42
6	1985.5	88.3	639	4	US-09-509-814A-4
7	1985.5	88.2	640	4	US-09-509-814A-8
8	1994.5	88.1	640	4	US-09-509-814A-6
9	1948.5	86.1	639	4	US-09-509-814A-1
10	1948.5	86.1	640	4	US-09-509-814A-2
11	1514	66.9	345	4	US-09-512-251A-10
12	1514	66.9	345	4	US-09-515-150A-10
13	1514	66.9	345	4	US-09-196-281-13
14	443	19.6	659	3	US-08-894-818B-1
15	443	19.6	659	3	US-08-894-818B-1
16	443	19.6	659	3	US-09-445-472-12
17	443	19.6	659	3	US-09-445-472-1
18	443	19.6	659	3	US-08-894-818B-3
19	443	19.6	659	3	US-09-445-472-4
20	443	19.6	659	3	US-08-894-818B-35
21	443	19.6	659	3	US-09-445-472-16
22	443	19.6	659	3	US-08-894-818B-5
23	443	19.6	659	3	US-09-000-016-7
24	443	19.6	659	3	US-09-514-340-7
25	443	19.6	659	3	US-09-000-016-4
26	443	19.6	659	3	US-09-514-340-4
27	443	19.6	659	3	US-09-000-016-2

28	305	13.5	903	1	US-08-750-532-1	Sequence 1, Appl1
29	305	13.5	1398	1	US-08-750-532-9	Sequence 8, Appl1
30	305	13.5	1398	4	US-08-894-818B-8	Sequence 9, Appl1
31	305	13.5	1398	4	US-09-445-472-6	Sequence 6, Appl1
32	280.5	12.4	237	1	US-08-750-532-18	Sequence 18, Appl1
33	272	12.0	275	2	US-08-750-406A-1	Sequence 1, Appl1
34	272	12.0	275	3	US-09-327-118-1	Sequence 1, Appl1
35	269	11.9	269	1	US-08-431-387-5	Sequence 5, Appl1
36	269	11.9	269	1	US-08-431-387-6	Sequence 6, Appl1
37	269	11.9	269	1	US-08-322-677A-10	Sequence 10, Appl1
38	269	11.9	269	1	US-08-322-676-10	Sequence 10, Appl1
39	269	11.9	269	2	US-08-140-083A-10	Sequence 10, Appl1
40	269	11.9	269	3	US-08-898-218-10	Sequence 10, Appl1
41	269	11.9	269	3	US-08-898-793-10	Sequence 10, Appl1
42	269	11.9	269	3	US-09-255-502-5	Sequence 5, Appl1
43	269	11.9	269	3	US-09-024-532-3	Sequence 3, Appl1
44	269	11.9	269	3	US-08-269-050-4	Sequence 4, Appl1
45	269	11.9	269	3	US-08-090-207-1	Sequence 1, Appl1

ALIGNMENTS

```
RESULT 1
US-09-104-623A-4
; Sequence 4, Application US/09104623A
; Patent No. 6303752
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Patum, Tine Muxoll
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Roggen, Erwin Ludo
; TITLE OF INVENTION: A Modified Polypeptide
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESS: No. 6303752 No. 6303752disk of No. 6303752th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS Windows Version 2.0
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,623A
; FILING DATE: 25-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 5256, 200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; STRAIN: Bacillus sp. Y
; US-09-104-623A-4
;
Query Match 98.9%; Score 2239; DB 4; Length 433;
Best Local Similarity 98.8%; Pred. No. 8.9e-165;
Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

QY 1 NDVARGIVKADVAQNNYGLVGGGVAVADTGTGRNDSMEHAFKITALYALRTN 60

```

Db      1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTRGRNDSWHEAFRGKITLALYALGRIN 60
Qy      61 NADDPNGHGTTHVAGSVLGNALNKGAPOANLVFQSIIMDSGGIGLPSNLNTLFSQAMNA 120
Db      61 NADDPNGHGTTHVAGSVLGNALNKGAPOANLVFQSIIMDSGGIGLPSNLNTLFSQAMNA 120
Qy      121 GARHTNSWGAAPNGATYANSROVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Db      121 GARHTNSWGAAPNGATYANSROVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Qy      181 VGATENRPSFSGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFLISARSSILAPSSFW 240
Db      181 VGATENRPSFSGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFLISARSSILAPSSFW 240
Qy      241 ANYSKXAYVGGTSMATPIVAGNVAOLREHFIKNRGITTPKPSLIKALLAGATDVGLGYP 300
Db      241 ANYSKXAYVGGTSMATPIVAGNVAOLREHFIKNRGITTPKPSLIKALLAGATDVGLGYP 300
Qy      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
Db      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
Qy      361 YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAQSGTYTIEVOAYNV 420
Db      361 YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAQSGTYTIEVOAYNV 420
Qy      421 PSGQRFSLAIYH 433
Db      421 PSGQRFSLAIYH 433

```

```

RESULT 2
US-09-019-532-4
; Sequence 4, Application US/09019532B
; Patent No. 6416756
; GENERAL INFORMATION:
; APPLICANT: Oleen, Anne Agerlin
; APPLICANT: Oleen, Anne Agerlin
; TITLE OF INVENTION: A Modified Enzyme for Skin Care
; FILE REFERENCE: 4922.204-US
; CURRENT APPLICATION NUMBER: US/09/019,532B
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: 0038/97
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 0754/97
; EARLIER FILING DATE: 1997-06-25
; EARLIER APPLICATION NUMBER: 60/051,381
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: PCT/DK98/00015
; EARLIER FILING DATE: 1998-01-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-019-532-4

```

```

Query Match      98.9%; Score 2239; DB 4; Length 433;
Best Local Similarity 98.8%; Pred. No. 8.9e-165;
Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy      1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTRGRNDSWHEAFRGKITLALYALGRIN 60
Db      1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTRGRNDSWHEAFRGKITLALYALGRIN 60
Qy      61 NADDPNGHGTTHVAGSVLGNALNKGAPOANLVFQSIIMDSGGIGLPSNLNTLFSQAMNA 120
Db      61 NADDPNGHGTTHVAGSVLGNALNKGAPOANLVFQSIIMDSGGIGLPSNLNTLFSQAMNA 120
Qy      121 GARHTNSWGAAPNGATYANSROVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Db      121 GARHTNSWGAAPNGATYANSROVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180

```

```

Db      121 GARHTNSWGAAPNGATYANSROVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Qy      181 VGATENRPSFSGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFLISARSSILAPSSFW 240
Db      181 VGATENRPSFSGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFLISARSSILAPSSFW 240
Qy      241 ANYSKXAYVGGTSMATPIVAGNVAOLREHFIKNRGITTPKPSLIKALLAGATDVGLGYP 300
Db      241 ANYSKXAYVGGTSMATPIVAGNVAOLREHFIKNRGITTPKPSLIKALLAGATDVGLGYP 300
Qy      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
Db      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
Qy      361 YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAQSGTYTIEVOAYNV 420
Db      361 YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAQSGTYTIEVOAYNV 420
Qy      421 PSGQRFSLAIYH 433
Db      421 PSGQRFSLAIYH 433

```

```

RESULT 3
US-09-338-746-4
; Sequence 4, Application US/09338746
; Patent No. 6638526
; GENERAL INFORMATION:
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Olsen, Arne A.
; APPLICANT: Fatum, Tine M.
; TITLE OF INVENTION: A Polypeptide-Polymer Conjugate
; FILE REFERENCE: 5619.200-US
; CURRENT APPLICATION NUMBER: US/09/338,746
; EARLIER FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: PA 1998 00809
; EARLIER FILING DATE: 1998-07-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-338-746-4

```

```

Query Match      98.9%; Score 2239; DB 4; Length 433;
Best Local Similarity 98.8%; Pred. No. 8.9e-165;
Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy      1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTRGRNDSWHEAFRGKITLALYALGRIN 60
Db      1 NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTRGRNDSWHEAFRGKITLALYALGRIN 60
Qy      61 NADDPNGHGTTHVAGSVLGNALNKGAPOANLVFQSIIMDSGGIGLPSNLNTLFSQAMNA 120
Db      61 NADDPNGHGTTHVAGSVLGNALNKGAPOANLVFQSIIMDSGGIGLPSNLNTLFSQAMNA 120
Qy      121 GARHTNSWGAAPNGATYANSROVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Db      121 GARHTNSWGAAPNGATYANSROVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Qy      181 VGATENRPSFSGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFLISARSSILAPSSFW 240
Db      181 VGATENRPSFSGSIADPNPHIAQFSSRGATRDGRIRKPDVTAPGTFLISARSSILAPSSFW 240
Qy      241 ANYSKXAYVGGTSMATPIVAGNVAOLREHFIKNRGITTPKPSLIKALLAGATDVGLGYP 300
Db      241 ANYSKXAYVGGTSMATPIVAGNVAOLREHFIKNRGITTPKPSLIKALLAGATDVGLGYP 300
Qy      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQAGKPLKISLWTDAPGSTTAS 360
Db      301 SGOQWGRVTLDSLNVAAYNEATALTGQKATYSFOAQAGKPLKISLWTDAPGSTTAS 360

```

DB 301 SGDQGWGRVTLDSLNVAAYVNEATLALATGOKATYSFOAQGRPLKSLVWTDAPGSTTAS 360
QY 361 YTLVNDLDLVITAPNGQKYVGNDFSYFDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
DB 361 YTLVNDLDLVITAPNGQKYVGNDFSYFDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
QY 421 PSGQRFSLAIHV 433
DB 421 PSGQRFSLAIHV 433
RESULT 4
US-08-873-479-43
Sequence 43, Application US/08873479
Patent No. 5891701
GENERAL INFORMATION:
APPLICANT: Sloma, Alan
TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
TITLE OF INVENTION: Having Protease Activity
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/873,479
FILING DATE: 12-JUN-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Agtis, Cheryl H
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 5251.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 635 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-873-479-43
Query Match 98.8%; Score 2239; DB 2; Length 635;
Best Local Similarity 98.8%; Pred. No. 1.5e-164;
Matches 428; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
DB 1 NDVARGIVKADVANNYGLYGQGVAVVADTGLDGRNDSVHEAFRGKITLALVAGRTN 60
DB 203 NDVARGIVKADVANNYGLYGQGVAVVADTGLDGRNDSVHEAFRGKITLALVAGRTN 262
QY 61 NADPENGHTHVAGSVLGNALNKGAPOANLVFQSIIMDSGGIGLPSNLNTLFSQAWNA 120
DB 263 NASDPNGHHTHVAGSVLGNALNKGAPOANLVFQSIIMDSGGIGLPSNLNTLFSQAWNA 322
QY 121 GARHINSNGAPVNGAYVTANSROVEYVNNMTVLFAAGNGSPNSGTSASAGTKAKAIT 180
DB 323 GARHINSNGAPVNGAYVTANSROVEYVNNMTVLFAAGNGSPNSGTSASAGTKAKAIT 382
QY 181 VGTATENYRPSFGSIADNPVHIAQFSSRGRATRDGRIRKPDVTAPGTFLSARSSLAPDSSFW 240
DB 383 VGTATENYRPSFGSIADNPVHIAQFSSRGRATRDGRIRKPDVTAPGTFLSARSSLAPDSSFW 442

QY 241 ANVSKAYMGCTSMATPIVAGVVAQUREHFICRGTTPKPSLIKALIAAGATDUGLGP 300
DB 443 ANVSKAYMGCTSMATPIVAGVVAQUREHFICRGTTPKPSLIKALIAAGATDUGLGP 502
QY 301 SGDQGWGRVTLDSLNVAAYVNEATLALATGOKATYSFOAQGRPLKSLVWTDAPGSTTAS 360
DB 503 SGDQGWGRVTLDSLNVAAYVNEATLALATGOKATYSFOAQGRPLKSLVWTDAPGSTTAS 562
QY 361 YTLVNDLDLVITAPNGQKYVGNDFSYFDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
DB 563 YTLVNDLDLVITAPNGQKYVGNDFSYFDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 622
QY 421 PSGQRFSLAIHV 433
DB 623 PSGQRFSLAIHV 635
RESULT 5
US-08-873-479-42
Sequence 42, Application US/08873479
Patent No. 5891701
GENERAL INFORMATION:
APPLICANT: Sloma, Alan
TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
TITLE OF INVENTION: Having Protease Activity
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/873,479
FILING DATE: 12-JUN-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Agtis, Cheryl H
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 5251.000-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 641 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-873-479-42
Query Match 90.1%; Score 2040; DB 2; Length 641;
Best Local Similarity 88.9%; Pred. No. 3.4e-149;
Matches 385; Conservative 24; Mismatches 24; Indels 0; Gaps 0;
DB 1 NDVARGIVKADVANNYGLYGQGVAVVADTGLDGRNDSVHEAFRGKITLALVAGRTN 60
DB 209 NDVARGIVKADVANNYGLYGQGVAVVADTGLDGRNDSVHEAFRGKITLALVAGRTN 268
QY 61 NADPENGHTHVAGSVLGNALNKGAPOANLVFQSIIMDSGGIGLPSNLNTLFSQAWNA 120
DB 269 NADPENGHTHVAGSVLGNALNKGAPOANLVFQSIIMDSGGIGLPSNLNTLFSQAWNA 328

QY 121 GARIHNSWGAIPYNGAYTANSROVDEYERNNDMTVLFAAGNBPNSGTTISAPGTAKNAIT 180
DB 329 GARIHNSWGAIPYNGAYTTDSRVYDDYKNDMTLLFAAGNBPNSGTTISAPGTAKNAIT 388
QY 181 VGATENYRPSFGSIADNPNIHIAOFSSRGATROGRIPDYATPOTFILSARSLADDSF 240
DB 389 VGATENYRPSFGSIADNPNIHIAOFSSRGATROGRIPDYATPOTFILSARSLADDSF 448
QY 241 ANNSKYAVGGSMTPIVAGNVAOLREHFIKNGITPKPSLIXKALJAGATDVLGP 300
DB 449 ANHDSKYAVGGSMTPIVAGNVAOLREHFIKNGITPKPSLIXKALJAGATDVLGP 508
QY 301 SGDQWGRVTLDKSLNVAAYNEATLTTGOKATYSFOTQAGKPKLSLWTDAPGSTTAS 360
DB 509 NGNQGWRVTLDKSLNVAAYNEATLTTGOKATYSFOTQAGKPKLSLWTDAPGSTTAS 568
QY 361 YTLVNDLVLITPNOCKYVNDPNDNNMDGRNNVNTLNPQSTYITIEVOAYN 420
DB 569 LTLVNDLVLITPNOCKYVNDPNDNNMDGRNNVNTLNPQSTYITIEVOAYN 628
QY 421 PPSGQRFSLAIYH 433
DB 629 PPSGQRFSLAIYH 641

RESULT 6

US-09-509-814A-4
; Sequence 4, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKETI, KATSUSHISA
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-4

Query Match 88.3%; Score 1998.5; DB 4; Length 639;
Best Local Similarity 88.0%; Pred. No. 5.3e-146;
Matches 382; Conservative 26; Mismatches 25; Indels 1; Gaps 1;
QY 1 NDVARGIVKADVQNNYGLYGQGOVAVADTGLDTRNDSMHEAFRGKITLALYALGRIN 60
DB 206 NDVARGIVKADVQNNYGLYGQGOVAVADTGLDTRNDSMHEAFRGKITLALYALGRIN 265
QY 61 NANDPNHGHTHVAGSVLGN-ALNKGAPOANLVFOSIMSSGGLGGLPSNLNTLFSQAWN 119
DB 266 NANDPNHGHTHVAGSVLGN-ALNKGAPOANLVFOSIMSSGGLGGLPSNLNTLFSQAWN 325
QY 120 AGARHNSWGAIPYNGAYTANSROVDEYERNNDMTVLFAAGNBPNSGTTISAPGTAKNAI 179
DB 326 AGARHNSWGAIPYNGAYTANSROVDEYERNNDMTVLFAAGNBPNSGTTISAPGTAKNAI 385
QY 180 TVGATENYRPSFGSIADNPNIHIAOFSSRGATROGRIPDYATPOTFILSARSLADDSF 239

DB 386 TVGATENYRPSFGSIADNPNIHIAOFSSRGATROGRIPDYATPOTFILSARSLADDSF 445
QY 240 ANNSKYAVGGSMTPIVAGNVAOLREHFIKNGITPKPSLIXKALJAGATDVLGP 299
DB 446 ANHDSKYAVGGSMTPIVAGNVAOLREHFIKNGITPKPSLIXKALJAGATDVLGP 505
QY 300 SGDQWGRVTLDKSLNVAAYNEATLTTGOKATYSFOTQAGKPKLSLWTDAPGSTTAS 359
DB 506 NGNQGWRVTLDKSLNVAAYNEATLTTGOKATYSFOTQAGKPKLSLWTDAPGSTTAS 565
QY 360 SYTLVNDLVLITPNOCKYVNDPNDNNMDGRNNVNTLNPQSTYITIEVOAYN 419
DB 566 SYTLVNDLVLITPNOCKYVNDPNDNNMDGRNNVNTLNPQSTYITIEVOAYN 625
QY 420 PPSGQRFSLAIYH 433
DB 626 PPSGQRFSLAIYH 639

RESULT 7

US-09-509-814A-8
; Sequence 8, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAKETI, KATSUSHISA
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509, 814A
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-8

Query Match 88.2%; Score 1995.5; DB 4; Length 640;
Best Local Similarity 87.6%; Pred. No. 9.1e-146;
Matches 380; Conservative 28; Mismatches 25; Indels 1; Gaps 1;
QY 1 NDVARGIVKADVQNNYGLYGQGOVAVADTGLDTRNDSMHEAFRGKITLALYALGRIN 60
DB 207 NDVARGIVKADVQNNYGLYGQGOVAVADTGLDTRNDSMHEAFRGKITLALYALGRIN 266
QY 61 NANDPNHGHTHVAGSVLGN-ALNKGAPOANLVFOSIMSSGGLGGLPSNLNTLFSQAWN 119
DB 267 NANDPNHGHTHVAGSVLGN-ALNKGAPOANLVFOSIMSSGGLGGLPSNLNTLFSQAWN 326
QY 120 AGARHNSWGAIPYNGAYTANSROVDEYERNNDMTVLFAAGNBPNSGTTISAPGTAKNAI 179
DB 327 AGARHNSWGAIPYNGAYTANSROVDEYERNNDMTVLFAAGNBPNSGTTISAPGTAKNAI 386
QY 180 TVGATENYRPSFGSIADNPNIHIAOFSSRGATROGRIPDYATPOTFILSARSLADDSF 239
DB 387 TVGATENYRPSFGSIADNPNIHIAOFSSRGATROGRIPDYATPOTFILSARSLADDSF 446
QY 240 ANNSKYAVGGSMTPIVAGNVAOLREHFIKNGITPKPSLIXKALJAGATDVLGP 299
DB 447 ANHDSKYAVGGSMTPIVAGNVAOLREHFIKNGITPKPSLIXKALJAGATDVLGP 506
QY 300 SGDQWGRVTLDKSLNVAAYNEATLTTGOKATYSFOTQAGKPKLSLWTDAPGSTTAS 359

```
Db 507 PNGNGQWGRVLTLDKSLNVAAYVNESSLSLTSQAKTYSFATACKPKISLVSDAPASTTA 566
QY 360 STTLVNDLDLVITAPNGQKTVGNDSYPYDNNWDRNNVENVFINAPOSQGYTTEVQAYN 419
Db 567 SYTLVNDLDLVITAPNGTQYVGNDFTSYDNDWDRNNVENVFINAPOSQGYTTEVQAYN 626
QY 420 VPSGQRFSLATVH 433
Db 627 VPVGPQFSLATVH 640

RESULT 8
US-09-509-814A-6
; Sequence 6, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1997-06-08
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-6

Query Match 88.1%; Score 1994.5; DB 4; Length 640;
Best Local Similarity 87.6%; Pred. No. 1.1e-145;
Matches 380; Conservative 28; Mismatches 25; Indels 1; Gaps 1;

QY 1 NNVAGIVADVAONNNGLYGGQVVAADTGLDGRNDSMEHAFRGKITALVALGRIN 60
Db 207 NNVAGIVADVAQSSVGLYGGQVVAADTGLDGRNDSMEHAFRGKITALVALGRIN 266
QY 61 NNDPNHGSHVAGSVLGN-ALNKGMAPQANLVFOSINDSSGGJGGLPSNINTLFSQAKN 119
Db 267 NNDPNHGSHVAGSVLGNSTNKGMAPQANLVFOSINDSSGGJGGLPSNINTLFSQAKN 326
QY 120 AGARHTNSWGAIPVNGAYTANSROVDEYVRNNDMTVLPAAGNEGNSGTTIAPGTAKAI 179
Db 327 AGARHTNSWGAIPVNGAYTANSROVDEYVRNNDMTVLPAAGNEGNSGTTIAPGTAKAI 386
QY 180 TVGATEENYRPSFGSIADNNPHIAQSSRGATPDGRIPKDVTAAPGFIIISASSSLAPDSF 239
Db 387 TVGATEENYRPSFGSIADNNPHIAQSSRGATPDGRIPKDVTAAPGFIIISASSSLAPDSF 446
QY 240 WANYSKVAVMGTSMAPIVAGNVAQJREHPIKRGITPKESLIKAALIGADTVGIGY 299
Db 447 WANYSKVAVMGTSMAPIVAGNVAQJREHPIKRGITPKESLIKAALIGADTVGIGY 506
QY 300 PSGDQMGWVTLDKSLNVAAYNEATALTGOKATYSFOTOAGKPKIXISLVMTDAPGSTTA 359
Db 507 PNGNGQWGRVLTLDKSLNVAAYNESSLSLTSQAKTYSFATACKPKISLVSDAPASTTA 566
QY 360 SYTLVNDLDLVITAPNGQKTVGNDSYPYDNNWDRNNVENVFINAPOSQGYTTEVQAYN 419
Db 567 SYTLVNDLDLVITAPNGTQYVGNDFTSYDNDWDRNNVENVFINAPOSQGYTTEVQAYN 626
```

```
QY 420 VPSGQRFSLATVH 433
Db 627 VPVGPQFSLATVH 640

RESULT 9
US-09-509-814A-1
; Sequence 1, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1997-06-08
; PRIOR APPLICATION NUMBER: JP 9-274570
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (45)..(45)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (70)..(70)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (89)..(89)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (102)..(102)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (105)..(105)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (130)..(130)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
```

LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (359)..(359)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (594)..(594)

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid

US-09-509-814A-1

Query Match 86.1%; Score 1948.5; DB 4; Length 639;
Best Local Similarity 86.4%; Pred No. 3.8e-142;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLYGGQVVAVADTGLDGRNDSMHEAFRGKITALYALGRN 60
DB NDVARGIVKADVAQSSVGLYGGQIVAVADTGLDGRNDSMHEAFRGKITALYALGRN 265
QY 61 NADPNRGHTVAGSVYGN-ALNKGAPOANTVFOGIMSSGGLGSPNLTLESOAMN 119
DB NADPNRGHTVAGSVYGNKXNKGNAPQANLVFOGIMSSGGLGSPNLTLESOAMN 325
QY 120 AGARHTNSWGA PYNGAYTANSRQVDEYVANNMTVLFAANGEPNSGTISAPGTAKNAI 179
DB AGARHTNSWGA AVNGAYTANSDRVDDYAKNMTILFAANGEXPNGGTISAPGTAKNAI 385
QY 180 TVGATENYRPSFGSIANPNHIAQFSRGATRGRIKPDVTAGTFLSRSSSLADDSF 229
DB TVGATENYRPSFGSIANPNHIAQFSRGATRGRIKPDVTAGTFLSRSSSLADDSF 445
QY 240 WANTSXYAVWGTSVATPIVAGVVAQREHFTKNGITRPSLILAAALTAGADVGLGY 299
DB WANTSXYAVWGTSVATPIVAGVVAQREHFTKNGITRPSLILAAALTAGADVGLGY 505
QY 300 PSCDQWGRVTLDKSLNVAAYVNEATLTTQKATYSFOIQAQAPLISLWMTAPGSTTA 359
DB PSCDQWGRVTLDKSLNVAAYVNEATLTTQKATYSFOIQAQAPLISLWMTAPGSTTA 565
QY 506 PNGQGWGRVTLDKSLNVAAYVNEATLTTQKATYSFOIQAQAPLISLWMTAPGSTTA 565
DB PNGQGWGRVTLDKSLNVAAYVNEATLTTQKATYSFOIQAQAPLISLWMTAPGSTTA 625
QY 420 VPSPGQRFSLAIYH 433
DB VPSPGQRFSLAIYH 639

RESULT 10
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION: MIKIO
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUYOSHI
APPLICANT: SAEKI, KATSUHISA
APPLICANT: KUBOTA, HIROKI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEIWA, YASUSHI
APPLICANT: SHIKATA, SHITSUM
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509, 814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08


```
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid
```

```
NAME/KEY: misc_feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-2
```

Query Match 86.1%; Score 1948.5; DB 4; Length 640;
Best Local Similarity 86.4%; Pred. No. 3.8e-142;
Matches 375; Conservative 20; Mismatches 38; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNTGVGGGVAVADTGLDTGNDSSMGEAFPGKLTALYAGRTN 60
DB 207 NDVARGIVKADVAQNTGVGGGVAVADTGLDTGNDSSMGEAFPGKLTALYAGRTN 266

```

QY 6 NADPNHGCHGVAVGVLGN-ALNKGAPOANI VPOSIMDSOGCLOGLPSENLTIFSOAMN 119
Db 267 NADPNTHGHVAVGVLGNKXTKGNAPOLNVPOSIMDSXGGJGGLPSNLQTLFSSQAS 326
QY 120 AGARHHTNSGAEVNGAYTANRSKOVDEYRRNMDVLFAAGNEGNSGTISAPGTAGNAI 179
Db 327 AGARHHTNSGAEVNGAYTTLDSHNVDDYRKNDMTLFLAAGNEXNGTISAPTAGNAI 386
QY 180 TVGATENTYRRSFGSIALDNPNHILQSSSKALTRDKRIKDYVAPGTFTILSARSLAPDSF 239
Db 387 TVGATENTLRPSFGSYADNIHVAQFSSRPTKDKRIKEDVAPGTXILSARSLAPDSF 446
QY 240 WAYSNSKAAVGGGSMATPIVAGNVQALEHRIKRGITPRSLIKALLIAGADVGLGY 229
Db 447 WANDHSKTAIMGSTSMATPIVAGNVQALEHREVKRGITPRSLIKALLIAGADKSLGY 506
QY 300 PSGDQNGRVTLDKSLNVAVNEALATLTGQKATYSFOTQKGPLKISLWTDAPGSTTA 359
Db 507 PNGQNGRVTLDKSLNVAVNESSKLSLSQQAQTLXFLATNGKPLKISLVMSDPASTTA 566
QY 360 SYTLVNDLDELVTAPNGQKITVGNDSYPRDNNWGRNNVENVFTINAPSGTYTILEVQAYN 419
Db 567 SYTLVNDLDELVTAPNGTYVGNDEFXPKXKNWDRNNVENVFTINXPSGTYTILEVQAYN 626
QY 420 VPSGQRFSLATVH 433
Db 627 VPSGQRFSLATVH 640

```

```

RESULT 11
US-09-512-251A-10
; Sequence 10, Application US/09512251A
; Patent No. 6555355
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mixkeisen, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5349, 204-US
; CURRENT APPLICATION NUMBER: US/09/512,251A
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 345
; TYPE: FRT
; ORGANISM: Bacillus
; US-09-512-251A-10

```

	Query Match	66.9%	Score 1514	DB 4	Length 345
	Best Local Similarity	90.2%	Pred. No. 52=109		
	Matches 286	Conservative 18	Mismatches 13	Indels 0	Gaps 0
QY	1	NDVARGIVADVAQNNTYGLVGGGGVVAVADGLDTRNDSSMHEAFEGKITALYALGRTN	60		
Db	29	NDVARGIVADVAQNNTYGLVGGGGVVAVADGLDTRNDSSMHEAFEGKITALYALGRTN	88		
QY	61	NANDPFGHGTAVAGSVLGNALNKGAPOANLFGOSIMSSGGAGGLPSNNTLFSQAWNA	120		
Db	89	NANDPFGHGTAVAGSVLGNATNKGAPOANLFGOSIMSSGGAGGLPSNNTLFSQAWNA	148		
QY	121	GARIHNSMGAPINNGAYITANSRQVEYKRNNDMTYLPAGNEGPNSGTISAPGTAAAT	180		
Db	149	GARIHNSMGAPINNGAYITANSRQVDDYKRNNDMTYLPAGNEGSGTISAPGTAAAT	208		
QY	181	VGATENYRPSFSLDNPWHIAQFSRSRATDGRIKEPDTAPGFTILSARSLAPDSSFW	240		
Db	209	VGATENYRPSFSSYADNINHWAFSSRPTDGRIKEDYVAPGYILISAASSLAPDSSFW	268		
QY	241	ANVNSKYAAVMGTSMAPIVAGNAQLAEHETIKRGITTPRSLIKALLAGADVGLGYP	300		
Db	269	ANVNSKYAAMGTSMAPIVAGNAQLAEHETIKRGVATPPRSLIKALLAGADVGLGYP	328		

```
QY      301  SGDQGMGRVTLDKSLNV 317
          :|:|||||
Db      329  NGNQGMGRVTLDKSLNV 345
```

RESULT 12

US-09-515-150A-10
; Sequence 10, Application US/09515150A
; Patent No. 6558938
; Inventor: INVENTOR

APPLICANT: Hansen, Peter
 APPLICANT: Bauditz, Peter
 APPLICANT: Mikkelson, Frank
 APPLICANT: Andersen, Kim
 TITLE OF INVENTION: Protease Variants and Compositions
 FILE REFERENCE: 5348, 204-US
 CURRENT APPLICATION NUMBER: US/09/515,150A
 CURRENT FILING DATE: 2000-02-29
 NUMBER OF SEQ. ID NOS: 12
 SOFTWARE: PatentIn version 3.1
 SEQ. ID NO 10
 LENGTH: 345
 TYPE: PRT
 ORGANISM: Bacillus
 US-09-515-150A-10

Query Match	66.9%	Score 1514;	DB 4;	Length 345;
Best Local Similarity	90.2%	Pred. No. 5.2e-109;		
Matches 286;	Conservative 18;	Mismatches 13;	Indels 0;	Gaps 0;

```

QY      1NNVAGGIYKADVAQNNNGLYXQGGQVAVVADTGLDTRDNRNSMHEAFRGKITLTVLAGSTN
Db      23NNVAGGIYKADVAQNNNGFLLYGGQIYAVADTGLDTRDNRNSMHEAFRGKITLTVLAGSTN
QY      61NANDENGHGHTVAGSVLGNALLNKGMAPOANLVFQSIIMDSGGI/GGLPSNINLTVLFSQAMNA
Db      89NANDENGHGHTVAGSVLGNATNKGMAPOANLVFQSIIMDSGGI/GGLPAVLQTLFQAVSA
QY      121GARHTNSWGA PVVAGATVANSRQVDEVYRRNNDMTVLFAQNGNGPNSGTSAPGTAKNAIT
Db      149GARHTNSWGA PVVAGATVNTDSRVNDDYVRKNDMTLLFAAGNGPGSGTISABGTAKNAIT
QY      181VGATENYRPSGSIADPNPHIAQFSSRGATRDGRIPDYTAAGFTLLSARSLAPDSSFW
Db      209VGATENYRPSGSIADPNHVAQFSSRGATRDGRIPDYTAAGFTLLSARSLAPDSSFW
QY      241ANYNKKYVMGCTSMAPPIYAGNVAQIREFHFNRRGTPKSLIKALAIAGATDVGLGYP
Db      265ANHDSKYVMGCTSMAPPIYAGNVAQIREFHFNRRGTPKSLIKALAIAGATDVGLGYP
QY      301SGDQGMGRVTLDKSLNV 317
Db      329NGNGMGGRVTLDKSLNV 345

RESULT 13
US-09-196-281-13
/ Sequence 13, Application US/09196281A
/ Patent No. 6605458
/ GENERAL INFORMATION:
/ APPLICANT: Hansen, Peter K.
/ APPLICANT: Baudeitz, Peter
/ APPLICANT: Mikkelsen, Frank
/ TITLE OF INVENTION: Protease Variants And Compositions
/ FILE REFERENCE: 5435.200-US
/ CURRENT APPLICATION NUMBER: US/09/196,281A
/ EARLIER FILING DATE: 1998-11-19
/ EARLIER APPLICATION NUMBER: 1332/97
/ NUMBER OF SEQ ID NOS: 18
/ SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 13

```

```

; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-09-196-281-13
Query Match
Best Local Similarity 66.9%; Score 1514; DB 4; Length 345;
Best Local Similarity 90.2%; Pred. No. 5.2e-109;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIKADYKQNNYGLYGGQVVAADTGLDTRGRDSSMHEAFRKTITALLY-ALGRTNNANDPN 60
DB 29 NDVARGIKADYKQNNYGLYGGQVVAADTGLDTRGRDSSMHEAFRKTITALLYALGRTN 88
QY 61 NANGENGHTHVAGSVLGNALNKGMAPQANLVFQSGIMDSGGGLGSPNLNTLFSQAWNA 120
DB 89 NANGENGHTHVAGSVLGNALNKGMAPQANLVFQSGIMDSGGGLGSPNLNTLFSQAWNA 148
QY 121 GARHTNSWGAQVNGAYTANRQVDEYRANNDMTVLPAAAGNEGPNSGTISAPGTAKNAIT 180
DB 149 GARHTNSWGAQVNGAYTANRQVDEYRANNDMTVLPAAAGNEGPNSGTISAPGTAKNAIT 208
QY 181 VGATENYRPSFGSIADNPNHIAQSSRGATRDGR:KPYTAAGTFTLSARSLAPDSSEW 240
DB 209 VQATENYRPSFGSIADNPNHIAQSSRGATRDGR:KPYTAAGTFTLSARSLAPDSSEW 268
QY 241 ANYNSKRYAMGTSMTPTIVAGNVAQLREHFIKNRGITPKPSLIIKALIIAGATDVGLGYP 300
DB 269 ANHDSKYAMGTSMTPTIVAGNVAQLREHFIKNRGITPKPSLIIKALIIAGATDVGLGYP 328
QY 301 SGDOGWRVTLDKSLNV 317
DB 329 NGNQGWRVTLDKSLNV 345

RESULT 14
US-08-894-818B-1
; Sequence 1, Application US/08894818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masamori
; APPLICANT: ASADA, Kiyozo
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-MAY-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: PCT/JP96/03253
; APPLICATION NUMBER: 07-NOV-1996
; PRIOR APPLICATION DATA: JP 322285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
```

```

; REFERENCE/DOCKET NUMBER: TAKAKURA-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-818B-1
Query Match
Best Local Similarity 19.6%; Score 443; DB 3; Length 659;
Best Local Similarity 31.1%; Pred. No. 4.8e-26;
Matches 137; Conservative 66; Mismatches 170; Indels 68; Gaps 16;

QY 8 VKADVAKQNNYGLYGGQVVAADTGLDTRGRDSSMHEAFRKTITALLY-ALGRTNNANDPN 66
DB 145 IGADTYWNSLGYDSSGVVAIVDTGIDAN-----HPDLKGVITGVYDAVNGRSTPYDDQ 198
QY 67 GHGTHVAGSVLGNALNKGMAPQANLVFQSGIM--DSGGGLGSPNLNTLFSQAWNA 120
DB 199 GHGTHVAGSVLGNALNKGMAPQANLVFQSGIM--DSGGGLGSPNLNTLFSQAWNA 258
QY 121 GARHTNSWGAQVNGAYTANRQVDEYRANNDMTVLPAAAGNEGPNSGTISAPGTAKNAIT 180
DB 259 GIRVIMLSLSSQSSGDTSLSQAVNNAMDAQIVCVAAGNSGPNYTVGSPAAASKVIT 318
QY 181 VGATENYRPSFGSIADNPNHIAQSSRGATRDGR:KPYTAAGTFTLSARSLAPDSSEW 240
DB 319 VQA-----VSDNIIASFSSRGPTADGRLEPYVAAROVDIARAS---GTSNG 364
QY 241 ANYNSKRYAMGTSMTPTIVAGNVAQLREHFIKNRGITP--KPSLIIKALIIAGATDVG- 296
DB 365 TPINDYTTKASGTSMTPTIVAGNVAQLREHFIKNRGITP--KPSLIIKALIIAGATDVG- 421
QY 297 LGYPSDQGWGRVTLDKSLNVAVNNAETALTGQ-----KATYSPQTQKPKISLVMT 351
DB 422 IAY-----GAKRVNVTKA--IKTDYAKLFTGSVADKSAHTTFVSGATFTATLYND 474
QY 352 DAGSTTASYTLVNDLVTAPNGQKVGNDPSYEDNNMGRRNVENVFINAPQSGTY 411
DB 475 -----TGSSDIDLVDYDN-----GNEVDYGYTAYY-----GFEKVGYNPFRAGT 515
QY 412 TIEVQAVNPSGQPSLSIAIV 432
DB 516 TVKVVSY--KGANYQVDVY 533

RESULT 15
US-09-445-472-12
; Sequence 12, Application US/09445472
; Patent No. 638726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOTO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/09/445,472
; CURRENT FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-09-445-472-12
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:09:09 ; Search time 33.1412 Seconds

(without alignments)
3418.697 Million cell updates/sec

Title: US-09-985-689A-3

Perfect score: 2263

Sequence: 1 NDVARGIVKADVADVAQNNYGLY.....EYQATNVDSGGFQRFSLATVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 26161801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*

10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*

12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*

14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*

15: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*

16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2263	100.0	433	10	US-09-985-689A-3
2	2251	99.5	433	10	US-09-985-689A-5
3	2234	98.7	433	10	US-09-985-689A-4
4	2040	90.1	433	10	US-09-985-689A-7
5	2020.5	89.3	434	10	US-09-985-689A-6
6	2006.5	88.7	434	10	US-09-985-689A-2
7	1994.5	88.1	434	10	US-09-985-689A-1
8	1994.5	88.1	434	10	US-10-385-662-2
9	1514	66.9	345	14	US-10-385-662-2
10	1514	66.9	345	14	US-10-403-105-13
11	443	19.6	659	13	US-10-090-624-12
12	422.5	18.7	412	13	US-10-090-624-1
13	422.5	18.7	522	13	US-10-090-624-4
14	422.5	18.7	564	13	US-10-090-624-16
15	346.5	15.3	1237	14	US-10-314-657-4

16	345.5	15.3	1079	14	US-10-112-488-39	Sequence 39, Appl
17	329	14.5	1208	14	US-10-156-761-13251	Sequence 13251, A
18	327	14.4	1139	14	US-10-156-761-10856	Sequence 10856, A
19	305	13.5	1398	13	US-10-090-624-6	Sequence 6, Appl
20	296.5	13.1	580	10	US-09-927-827-55	Sequence 55, Appl
21	295	13.0	519	15	US-10-084-846A-114	Sequence 114, App
22	295	13.0	19725	15	US-10-084-846A-4	Sequence 4, Appl
23	282.5	12.5	1101	14	US-10-156-761-12934	Sequence 12934, A
24	281.5	12.4	368	12	US-10-344-231-3	Sequence 3, Appl
25	276	12.2	271	10	US-09-813-408-2	Sequence 2, Appl
26	271	12.0	271	14	US-10-242-549-56	Sequence 56, Appl
27	270	11.9	271	14	US-10-242-549-54	Sequence 54, Appl
28	270	11.9	271	14	US-10-242-549-60	Sequence 60, Appl
29	269	11.9	269	8	US-08-322-678-10	Sequence 10, Appl
30	269	11.9	269	9	US-09-837-635-16	Sequence 16, Appl
31	269	11.9	269	9	US-09-060-854B-6	Sequence 6, Appl
32	269	11.9	269	9	US-09-975-139-1	Sequence 1, Appl
33	269	11.9	269	9	US-09-976-114-8	Sequence 8, Appl
34	269	11.9	269	10	US-09-736-116-49	Sequence 49, Appl
35	269	11.9	269	13	US-10-075-907-1	Sequence 1, Appl
36	269	11.9	269	13	US-10-075-895-1	Sequence 1, Appl
37	269	11.9	269	14	US-10-033-325-6	Sequence 6, Appl
38	269	11.9	269	14	US-10-209-812-3	Sequence 3, Appl
39	269	11.9	269	14	US-10-104-693-4	Sequence 4, Appl
40	269	11.9	269	14	US-10-228-572-6	Sequence 6, Appl
41	269	11.9	269	14	US-10-324-152-5	Sequence 5, Appl
42	269	11.9	269	14	US-10-342-549-5	Sequence 5, Appl
43	269	11.9	269	14	US-10-313-853-1	Sequence 1, Appl
44	269	11.9	269	14	US-10-336-324-4	Sequence 4, Appl
45	269	11.9	269	14	US-10-336-324-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1

US-09-985-689A-3

Sequence 3, Application US/09985689A

Publication No. US20030022351A1

GENERAL INFORMATION:

APPLICANT: HATADA, YUJI

APPLICANT: OGAWA, AKINOBU

APPLICANT: KAGEYAMA, YASUSHI

APPLICANT: SATO, TOSYOSHI

APPLICANT: ARAKI, HIROYUKI

APPLICANT: SUMITOMO, NOBUYUKI

APPLICANT: OKUDA, MITSUYOSHI

APPLICANT: SAKETI, KATSUSHI

TITLE OF INVENTION: Alkaline proteases

FILE REFERENCE: 215483US0

CURRENT APPLICATION NUMBER: US/09/985,689A

PRIOR FILING DATE: 2002-07-01

PRIOR APPLICATION NUMBER: JP P2000-355166

PRIOR FILING DATE: 2000-11-22

PRIOR APPLICATION NUMBER: JP P2001-114048

NUMBER OF SEQ ID NOS: 7

SOFTWARE: PatentIn version 3.1

SEQ ID NO 3

LENGTH: 433

TYPE: PRT

ORGANISM: Bacillus sp.

US-09-985-689A-3

Query Match 100.0%; Score 2263; DB 10; Length 433;

Best Local Similarity 100.0%; Pred. No. 3, 5e-191;

Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVADVAQNNYGLVGQGVAVADTGDGTGRNDSMGEARFGKTTALYALGRTN 60

DB 1 NDVARGIVKADVADVAQNNYGLVGQGVAVADTGDGTGRNDSMGEARFGKTTALYALGRTN 60

OY 61 NADPNHGHTVAGSVLGNALNKGMAPQAVLVFQSIIMDSGGIGLPLSNITLTFQGAWNA 120

```

Db      61  NADDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGGLPSNLNTLFSQAMNA 120
Qy      121  GARIHTNSWGA PVNGAYTANSRQVDEVVRNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Db      121  GARIHTNSWGA PVNGAYTANSRQVDEVVRNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Qy      181  VGATENTRPSFGSIADPNPHIAQFSSRGATRDGRIPDVTAPGTFTLSARSSLAPOSSFW 240
Db      181  VGATENTRPSFGSIADPNPHIAQFSSRGATRDGRIPDVTAPGTFTLSARSSLAPOSSFW 240
Qy      241  ANYNSKYAYMGTSMA TPVAGNVAQLREHFINKRGITPSPSLIKALAGATDVGLGYP 300
Db      241  ANYNSKYAYMGTSMA TPVAGNVAQLREHFINKRGITPSPSLIKALAGATDVGLGYP 300
Qy      301  SGDOGMRVTLDSKLNAYVNEATALTGGKATYSFOQAQKPLKLSLWTDAPGSTTAS 360
Db      301  SGDOGMRVTLDSKLNAYVNEATALTGGKATYSFOQAQKPLKLSLWTDAPGSTTAS 360
Qy      361  YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
Db      361  YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
Qy      421  PSGPQRFSLAIHV 433
Db      421  PSGPQRFSLAIHV 433

```

RESULT 2

```

US-09-985-689A-5
; Sequence 5, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SAKO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483U0
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-5

```

```

Query Match      99.5%; Score 2251; DB 10; Length 433;
Best Local Similarity 99.3%; Pred. No. 4, 1e-190;
Matches 430; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy      1  NDVARGIVKADVQNNNYGLYGQGVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
Db      1  NDVARGIVKADVQNNNYGLYGQGVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
Qy      61  NADDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGGLPSNLNTLFSQAMNA 120
Db      61  NADDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGGLPSNLNTLFSQAMNA 120
Qy      121  GARIHTNSWGA PVNGAYTANSRQVDEVVRNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Db      121  GARIHTNSWGA PVNGAYTANSRQVDEVVRNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180

```

```

Qy      181  VGATENTRPSFGSIADPNPHIAQFSSRGATRDGRIPDVTAPGTFTLSARSSLAPOSSFW 240
Db      181  VGATENTRPSFGSIADPNPHIAQFSSRGATRDGRIPDVTAPGTFTLSARSSLAPOSSFW 240
Qy      241  ANYNSKYAYMGTSMA TPVAGNVAQLREHFINKRGITPSPSLIKALAGATDVGLGYP 300
Db      241  ANYNSKYAYMGTSMA TPVAGNVAQLREHFINKRGITPSPSLIKALAGATDVGLGYP 300
Qy      301  SGDOGMRVTLDSKLNAYVNEATALTGGKATYSFOQAQKPLKLSLWTDAPGSTTAS 360
Db      301  SGDOGMRVTLDSKLNAYVNEATALTGGKATYSFOQAQKPLKLSLWTDAPGSTTAS 360
Qy      361  YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
Db      361  YTLVNDLDLVITAPNGQKYVGNDFSYPDNNMDGRNNVENVFINAPQSGTYTIEVOAYNV 420
Qy      421  PSGPQRFSLAIHV 433
Db      421  PSGPQRFSLAIHV 433

```

RESULT 3

```

US-09-985-689A-4
; Sequence 4, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SAKO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483U0
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-4

```

```

Query Match      98.7%; Score 2234; DB 10; Length 433;
Best Local Similarity 98.6%; Pred. No. 1, 3e-188;
Matches 427; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Qy      1  NDVARGIVKADVQNNNYGLYGQGVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
Db      1  NDVARGIVKADVQNNNYGLYGQGVAVADTGLDTRGNDSSMHEAFRGKITLALYALGRTN 60
Qy      61  NADDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGGLPSNLNTLFSQAMNA 120
Db      61  NADDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGIGGLPSNLNTLFSQAMNA 120
Qy      121  GARIHTNSWGA PVNGAYTANSRQVDEVVRNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Db      121  GARIHTNSWGA PVNGAYTANSRQVDEVVRNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Qy      181  VGATENTRPSFGSIADPNPHIAQFSSRGATRDGRIPDVTAPGTFTLSARSSLAPOSSFW 240
Db      181  VGATENTRPSFGSIADPNPHIAQFSSRGATRDGRIPDVTAPGTFTLSARSSLAPOSSFW 240
Qy      241  ANYNSKYAYMGTSMA TPVAGNVAQLREHFINKRGITPSPSLIKALAGATDVGLGYP 300
Db      241  ANYNSKYAYMGTSMA TPVAGNVAQLREHFINKRGITPSPSLIKALAGATDVGLGYP 300

```

QY 301 SGDQMGRTYLDKSLNVAVNEATALTGOKATYSFOTQAGPKLISLWMTDAPGSTTAS 360
 Db 301 NGDQMGRTYLDKSLNVAVNEATALTGOKATYSFOTQAGPKLISLWMTDAPGSTTAS 360
 QY 361 YTLVNDLDELVTAPNGQKVVGNDFSYPDNNMDGNNVENVFINAPQSGTYTIEVOAYNV 420
 Db 361 YTLVNDLDELVTAPNGQKVVGNDFSYPDNNMDGNNVENVFINAPQSGTYTIEVOAYNV 420
 QY 421 PSGPQRFSLAIYH 433
 Db 421 PSGPQRFSLAIYH 433

RESULT 4
 US-09-985-689A-7
 ; Sequence 7, Application US/09985689A
 ; Publication No. US20030022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAWA, AKINORI
 ; APPLICANT: KAGEYAMA, YASUSHI
 ; APPLICANT: SATO, TSUYOSHI
 ; APPLICANT: ARAKI, HIROYUKI
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: SAEKI, KATSUHIISA
 ; TITLE OF INVENTION: Alkaline proteases
 ; FILE REFERENCE: 215483US0
 ; CURRENT APPLICATION NUMBER: US/09/985,689A
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: JP P2000-355166
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: JP P2001-114048
 ; PRIOR FILING DATE: 2001-04-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 433
 ; TYPE: PR1
 ; ORGANISM: Bacillus sp.
 US-09-985-689A-7

Query Match 90.1%; Score 2040; DB 10; Length 433;
 Best Local Similarity 88.9%; Pred. No. 1,8e-171;
 Matches 385; Conservative 24; Mismatches 24; Indels 0; Gaps 0;

QY 1 NDVARGIVADVAQNNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
 Db 1 NDVARGIVADVAQNNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
 QY 61 NANDNGHGTHTVAGSVLGNALNKGMAPQANLVFQSIMSSGGLGSPENLNTLFSQANNA 120
 Db 61 NANDNGHGTHTVAGSVLGNALNKGMAPQANLVFQSIMSSGGLGSPENLNTLFSQANNA 120
 QY 121 GARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 Db 121 GARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSIADNPVHIAQFSSRGATRDGRKRPVTAAGTITLSARSSLPADSSFW 240
 Db 181 VGATENYRPSFGSIADNPVHIAQFSSRGATRDGRKRPVTAAGTITLSARSSLPADSSFW 240
 QY 241 ANYSKYAYMGTSNATPIVAGNVAQLEHFIKRGITPKPSLAKALIAAGTIDVGLGYP 300
 Db 241 ANHDSKYAYMGTSNATPIVAGNVAQLEHFIKRGITPKPSLAKALIAAGTIDVGLGYP 300
 QY 301 SGDQMGRTYLDKSLNVAVNEATALTGOKATYSFOTQAGPKLISLWMTDAPGSTTAS 360
 Db 301 NGDQMGRTYLDKSLNVAVNEATALTGOKATYSFOTQAGPKLISLWMTDAPGSTTAS 360
 QY 361 YTLVNDLDELVTAPNGQKVVGNDFSYPDNNMDGNNVENVFINAPQSGTYTIEVOAYNV 420
 Db 361 YTLVNDLDELVTAPNGQKVVGNDFSYPDNNMDGNNVENVFINAPQSGTYTIEVOAYNV 420

Db 361 LTLVNDLDELVTAPNGKTYVGNDFTAQPDNNMDGNNVENVFINAPQSGTYTIEVOAYNV 420
 QY 421 PSGPQRFSLAIYH 433
 Db 421 VPSQTFSLAIYH 433

RESULT 5
 US-09-985-689A-6
 ; Sequence 6, Application US/09985689A
 ; Publication No. US20030022351A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HATADA, YUJI
 ; APPLICANT: OGAWA, AKINORI
 ; APPLICANT: KAGEYAMA, YASUSHI
 ; APPLICANT: SATO, TSUYOSHI
 ; APPLICANT: ARAKI, HIROYUKI
 ; APPLICANT: SUMITOMO, NOBUYUKI
 ; APPLICANT: SAEKI, KATSUHIISA
 ; TITLE OF INVENTION: Alkaline proteases
 ; FILE REFERENCE: 215483US0
 ; CURRENT APPLICATION NUMBER: US/09/985,689A
 ; CURRENT FILING DATE: 2002-07-01
 ; PRIOR APPLICATION NUMBER: JP P2000-355166
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: JP P2001-114048
 ; PRIOR FILING DATE: 2001-04-12
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 434
 ; TYPE: PR1
 ; ORGANISM: Bacillus sp.
 US-09-985-689A-6

Query Match 89.3%; Score 2020.5; DB 10; Length 434;
 Best Local Similarity 88.2%; Pred. No. 9.3e-170;
 Matches 383; Conservative 28; Mismatches 22; Indels 1; Gaps 1;

QY 1 NDVARGIVADVAQNNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
 Db 1 NDVARGIVADVAQNNYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
 QY 61 NANDNGHGTHTVAGSVLGNALNKGMAPQANLVFQSIMSSGGLGSPENLNTLFSQANNA 119
 Db 61 NANDNGHGTHTVAGSVLGNALNKGMAPQANLVFQSIMSSGGLGSPENLNTLFSQANNA 120
 QY 120 AGARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAI 179
 Db 120 AGARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAI 180
 QY 180 TVGATENYRPSFGSIADNPVHIAQFSSRGATRDGRKRPVTAAGTITLSARSSLPADSSFW 239
 Db 180 TVGATENYRPSFGSIADNPVHIAQFSSRGATRDGRKRPVTAAGTITLSARSSLPADSSFW 240
 QY 240 MANTSKYAYMGTSNATPIVAGNVAQLEHFIKRGITPKPSLAKALIAAGTIDVGLGYP 299
 Db 241 MANHDSKYAYMGTSNATPIVAGNVAQLEHFIKRGITPKPSLAKALIAAGTIDVGLGYP 300
 QY 300 PSGDQMGRTYLDKSLNVAVNEATALTGOKATYSFOTQAGPKLISLWMTDAPGSTTAS 359
 Db 301 PSGDQMGRTYLDKSLNVAVNEATALTGOKATYSFOTQAGPKLISLWMTDAPGSTTAS 360
 QY 360 SYTLVNDLDELVTAPNGQKVVGNDFSYPDNNMDGNNVENVFINAPQSGTYTIEVOAYNV 419
 Db 361 SYTLVNDLDELVTAPNGKTYVGNDFTAQPDNNMDGNNVENVFINAPQSGTYTIEVOAYNV 420
 QY 420 VPSQTFSLAIYH 433
 Db 421 VPSQTFSLAIYH 434

```

RESULT 6
US-09-985-689a-2
; Sequence 2, Application US/09985689a
; Publication No. US2003022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SAITO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUISHA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; PRIORITY FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689a-2

Query Match      88.7%; Score 2006.5; DB 10; Length 434;
Best Local Similarity 88.2%; Pred. No. 1.6e-168;
Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLYGQGVVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 60
QY 61 NADPNHGHTHVAGSVLGN ALNKGAPOANLVFQSIIMSSGGLGGLPSNLTLFSGQAN 119
DB 61 NADPNHGHTHVAGSVLGNATNKGAPQANLVFQSIIMSSGGLGGLPSNLTLFSGQAN 120
QY 120 AGARIHTNSGAPVNGAYTANSRQVDEYVANNMTVLFAAGNPGSGTISAPGAKNAI 179
DB 121 AGARIHTNSGAAVNGAYTTDSNVDDYRKNDMTILFAAGNPGSGTISAPGAKNAI 180
QY 180 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFILSARSSLADSSF 239
DB 181 TVGATENYRPSFGSYADNINHVAFQSSRGPTKGRIKPDVMAPGTYILSARSSLADSSF 240
QY 240 WANNYSKYAVMGSTMAPIVAGNVAQLREHFIKNGITPKPSLIKALIIAGATVGLGY 299
DB 241 WANNHDKYAVMGSTMAPIVAGNVAQLREHFIKNGITPKPSLIKALIIAGADVGLGY 300
QY 300 PSEGQGWGRVTLDKSLNVAAYNEATLTTGQKATYSFOTQAGKPKLISLWTDAPGSTTA 359
DB 301 PGNQGWGRVTLDKSLNVAAYNESSALSTSQKATYFTATAGKPKLISLWSDAPASTTA 360
QY 360 SYTLVNDLDTVITAPNGOKYVGNDFSYPDNNMDGNNVENVFINAPQSGTYTIEVOAYN 419
DB 361 SYTLVNDLDTVITAPNGRTYVGNDFSAFPDNNMDGNNVENVFINSQSGTYTIEVOAYN 420
QY 420 VPSPGQRFSLAIYH 433
DB 421 VPVGQPFSLAIYH 434

```

```

; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SAITO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUISHA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985, 689A
; PRIORITY FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689a-1

Query Match      88.1%; Score 1994.5; DB 10; Length 434;
Best Local Similarity 87.6%; Pred. No. 1.6e-167;
Matches 380; Conservative 28; Mismatches 25; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNYGLYGQGVVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 60
DB 1 NDVARGIVKADVAQSSYGLYGQGVVAVADTGLDTGRNDSMHEAFRGKITLALYALGRIN 60
QY 61 NADPNHGHTHVAGSVLGN ALNKGAPOANLVFQSIIMSSGGLGGLPSNLTLFSGQAN 119
DB 61 NADPNHGHTHVAGSVLGNSTNKGAPQANLVFQSIIMSSGGLGGLPSNLTLFSGQAN 120
QY 120 AGARIHTNSGAPVNGAYTANSRQVDEYVANNMTVLFAAGNPGSGTISAPGAKNAI 179
DB 121 AGARIHTNSGAAVNGAYTTDSNVDDYRKNDMTILFAAGNPGSGTISAPGAKNAI 180
QY 180 TVGATENYRPSFGSIADNPNHIAQFSSRGATRDGRIKPDVTAPGTFILSARSSLADSSF 239
DB 181 TVGATENYRPSFGSYADNINHVAFQSSRGPTKGRIKPDVMAPGTYILSARSSLADSSF 240
QY 240 WANNYSKYAVMGSTMAPIVAGNVAQLREHFIKNGITPKPSLIKALIIAGATVGLGY 299
DB 241 WANNHDKYAVMGSTMAPIVAGNVAQLREHFIKNGITPKPSLIKALIIAGADVGLGY 300
QY 300 PSEGQGWGRVTLDKSLNVAAYNEATLTTGQKATYSFOTQAGKPKLISLWTDAPGSTTA 359
DB 301 PGNQGWGRVTLDKSLNVAAYNESSALSTSQKATYFTATAGKPKLISLWSDAPASTTA 360
QY 360 SYTLVNDLDTVITAPNGOKYVGNDFSYPDNNMDGNNVENVFINAPQSGTYTIEVOAYN 419
DB 361 SYTLVNDLDTVITAPNGRTYVGNDFSPYDNNMDGNNVENVFINAPQSGTYTIEVOAYN 420
QY 420 VPSPGQRFSLAIYH 433
DB 421 VPVGQPFSLAIYH 434

```

```

RESULT 7
US-09-985-689a-1
; Sequence 1, Application US/09985689a
; Publication No. US2003022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI

```

```

RESULT 8
US-10-385-662-2
; Sequence 2, Application US/10385662
; Publication No. US20040002432A1
; GENERAL INFORMATION:
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAITO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: IZAWA, YOSHIFUMI
; APPLICANT: SAEKI, KATSUISHA
; APPLICANT: KOBAYASHI, TOHRU
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: Alkaline protease

```


Db 149 GARHTNSWGAFFVGTDDSRVDDYVRKNDWTLFAAGNPGSGTISAPGTAKNAIT 208
Qy 181 VGATENYRPSFGSIADNPENHIAOFSSRGATROGRIRKPDVTAGFTILSRSLADSSFW 240
Db 209 VGATENYRPSFGSIADNPENHIAOFSSRGATROGRIRKPDVTAGFTILSRSLADSSFW 268
Qy 241 ANYSKXAYWGTSMATPIVAGNVAQLREHFIKNGRITKPSLIXAALJAGATDYGLYP 300
Db 269 ANHDSKXAYWGTSMATPIVAGNVAQLREHFIKNGRITKPSLIXAALJAGATDYGLYP 328
Qy 301 SGDOGWRVTLDKSLNV 317
Db 329 NGNQGWRVTLDKSLNV 345

RESULT 11
US-10-090-624-12
; Sequence 12, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match 19.6%; Score 443; DB 13; Length 659;
Best Local Similarity 31.1%; Pred. No. 3,7e-30;
Matches 137; Conservative 66; Mismatches 170; Indels 68; Gaps 16;
Qy 8 VKADVAQNNGYGLGQGVAVADTGLDGRNDSKHEAFRGKITALY-ALGRTNNANPN 66
Db 145 TGAADYWNLSLGYSGVVAIVDTGIDAN----HPDLKGVITGWDVNGRSTPYDQ 198
Qy 67 GHGTHVAGSVLG---NALNKGAPOANLVFQSIW--DSSGGLGSPNLNTLFSQAWNA 120
Db 199 GHGTHVAGIAGTGVNSQYIGVAPGAKLVGVKVLADGSGSVSTIIADVMDVQNKXK 258
Qy 121 GAIHINSNCAFPYNGAVTANSRQVDEYRKNMDTYLFAAGNPGSGTISAPGTAKNAIT 180
Db 259 GILVIVLUSGSSQSGTDSLSQAVNNAWDAGLVVCAAGNSGPNYTVTGSFAASKYIT 318
Qy 181 VGATENYRPSFGSIADNPENHIAOFSSRGATROGRIRKPDVTAGFTILSRSLADSSFW 240
Db 319 VGA-----VDSNDINAFSSRGPADGRLEPEVVAQGVDIARAS---GTSNG 364
Qy 241 ANYSKXAYWGTSMATPIVAGNVAQLREHFIKNGRITP---KPSLIXAALJAGATDVG- 296
Db 365 TPINDYTTKASGTSMATPIVAGNVAQLREHFIKNGRITP---KPSLIXAALJAGATDVG- 421
Qy 297 LGYPSDQGWGRVTLDKSLNVAVNEATALTTCQ-----KATYSFOTQAGPLKISLVWT 351
Db 422 IAY-----GAGRNVYKKA--IKTDYAKLFTGSAVDKSAHHTFVSGATFVTATLYMD 474
Qy 352 DAFGSTTASTYTLVNDLDTVITAPNGQKYVGNDSFYDNNMDGRNNVENVFINAQSGTY 411
Db 475 -----TGSSDIDLVLYPD-----GNEVDYSYTYV---GFEKVGYYNFTAGTW 515

Qy 412 TIEQAYNVPSGPORFSLATV 432
Db 516 TVKVSY---KGAANYQVDVV 533

RESULT 12
US-10-090-624-1
; Sequence 1, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA-6
; CURRENT APPLICATION NUMBER: US/10/090,624
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-10-090-624-1

Query Match 18.7%; Score 422.5; DB 13; Length 412;
Best Local Similarity 31.4%; Pred. No. 1,2e-28;
Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;

Qy 16 NYGLYGGQVAVAVADTGLDGRNDSKHEAFRGKITALYALGRTNNAN-----DENGHG 69
Db 20 NLGYDGSGITIGITIDID-----ASHPDLQGV-----IGWVDVNGRSYDDHGHG 68
Qy 70 THVAGSVLG---NALNKGAPOANLVFQSIW--DSSGGLGSPNLNTLFSQAWNA 122
Db 69 THVASIAGTGAASNGKXKMAFGAKIAGIKVLGADGSGSITIIKGVAVADNKKYGI 128
Qy 123 RHHTNSWGAFFVGTANSRQVDEYRKNMDTYLFAAGNPGSGTISAPGTAKNAIT 182
Db 129 KVTLSLSSQSSSGTIDALSQAVNNAWDAGLVVVAAGNSGPNYTVTGSFAASKYIT 188
Qy 183 ATENYRPSFGSIADNPENHIAOFSSRGATROGRIRKPDVTAGFTILSRSLADSSFWAN 242
Db 189 AVDKY-----DVITSFSSRGPTADGRLEPEVVAQGVNIIAARAS---GTSNGQP 234
Qy 243 YNSKXAYWGTSMATPIVAGNVAQLREHFIKNGRITP---KPSLIXAALJAGATDVG- 298
Db 235 INDYTTAPGTSMATPIVAGNVAQLREHFIKNGRITP---KPSLIXAALJAGATDVG- 291
Qy 299 YPSGDQGWGRVTLDKSLNVAVNEATALTTCQ-----TYSFOTQAGPLKISLVWTD 353
Db 292 Y-----GAGRNVYKKA--YDNYAKLFTGSAVDKSAHHTFVSGATFVTATLYMD 344
Qy 354 PGSTASTYTLVNDLDTVITAPNGQKYVGNDSFYDNNMDGRNNVENVFINAQSGTY 413
Db 345 N-----SDIDLVLYPDNGNV---DYSY-----TAYYGFKVGYYNFTAGTW 385
Qy 414 EVQAYNVPSGPORFSLATV 432
Db 386 KVSYS---SGSANYQVDVV 401

RESULT 13
US-10-090-624-4
; Sequence 4, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:

APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/090,624
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 522
TYPE: PR
ORGANISM: Pyrococcus furiosus
FEATURE:
NAME/KEY: misc feature
LOCATION: (428)..(428)
OTHER INFORMATION: Xaa at position 428 is Gly or Val.
US-10-090-624-4

Query Match 18.7%; Score 422.5; DB 13; Length 522;
Best Local Similarity 31.4%; Pred. No. 1,7e-28;
Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;

QY 16 NYGLYGGQGVAVADTGLDGRNDSMEAFRGKITLALGRTNNAN-----DPNGHG 69
DB 20 NLGYDSGGITIGITIDGID-----ASHPDLQGV-----IGWDFVNGRSYPYDHHG 68
QY 70 THVAGSVLG-----NALNKGMAPQANLVFQSIM--DSSGGLGGLPSNLTLSQAMNAGA 122
DB 69 THVASIAAGTGAASNGKYGKAPGAKLAGIKVLGADSGSISITIKGVWAVDNKDKYGI 128
QY 123 RHITNSGAPVNGAYTANSRQVDEYVNNDMTVLFAAGNEGPNSGTISAPGTAKNATV 182
DB 129 KVINLSGSSGSDGTALSQAVNAAMDAGLVVVAAGNSGPNKTYTIGSPAASKYITVG 188
QY 183 ATENYRPSFGSIADNPNHIAQFSSRGATRDGRKPDVTAPGFTILSARSSLPDSSFWAN 242
DB 189 AVDKY-----DVITFSRGPETADGRKPEVVAAGNMIIARAS---GTSWGP 234
QY 243 YNSKYAVMGTSNATPIVAGNVAQLREHFIKNGITP---KPSLIKALIAGATDVG-LG 298
DB 235 INDYTTAPGTSNATPVAAGIALLLQ---AHSWTPDKVKTALIEADIVKDEIADIA 291
QY 299 YPSGDQGMGRVTLDKSLNVAVNEATALTTGOKA-----TYSFQTOAGKPLKISLVWTD 353
DB 292 Y-----GAGRNVAKAIN--YDNVAKLVFTGYVANKSGTHQFVIGASFVATLWMDNA 344
QY 354 PGSTTASYTLVNDLDTLTAFAVNGQKYVNDPSYPDNNDGRNNVENVFINAPQSGTYTI 413
DB 345 N-----SDLDLYLDPNGNQV---DYSY-----TAYYGFEXVGYNPTDGTWTI 385
QY 414 EVQAVNPSGQRFSLAI 432
DB 386 KVVSY---SGSANYQVDV 401

RESULT 14
US-10-090-624-16
Sequence 16, Application US/10090624
Publication No. US2002012335A1
GENERAL INFORMATION:
APPLICANT: TAKAKURA, Hikaru
APPLICANT: MORISHITA, Mio
APPLICANT: SHIMOJO, Tomoko
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE

FILE REFERENCE: TAKAKURA-6
CURRENT APPLICATION NUMBER: US/10/090,624
CURRENT FILING DATE: 2002-03-06
PRIOR APPLICATION NUMBER: 09/445,472
PRIOR FILING DATE: 1999-12-06
PRIOR APPLICATION NUMBER: 151969/1997
PRIOR FILING DATE: 1997-06-10
NUMBER OF SEQ ID NOS: 33
SOFTWARE: PatentIn version 3.0
SEQ ID NO 16
LENGTH: 654
TYPE: PR
ORGANISM: Pyrococcus furiosus
US-10-090-624-16

Query Match 18.7%; Score 422.5; DB 13; Length 654;
Best Local Similarity 31.4%; Pred. No. 2,4e-28;
Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;

QY 16 NYGLYGGQGVAVADTGLDGRNDSMEAFRGKITLALGRTNNAN-----DPNGHG 69
DB 152 NLGYDSGGITIGITIDGID-----ASHPDLQGV-----IGWDFVNGRSYPYDHHG 200
QY 70 THVAGSVLG-----NALNKGMAPQANLVFQSIM--DSSGGLGGLPSNLTLSQAMNAGA 122
DB 201 THVASIAAGTGAASNGKYGKAPGAKLAGIKVLGADSGSISITIKGVWAVDNKDKYGI 260
QY 123 RHITNSGAPVNGAYTANSRQVDEYVNNDMTVLFAAGNEGPNSGTISAPGTAKNATV 182
DB 261 KVINLSGSSGSDGTALSQAVNAAMDAGLVVVAAGNSGPNKTYTIGSPAASKYITVG 320
QY 183 ATENYRPSFGSIADNPNHIAQFSSRGATRDGRKPDVTAPGFTILSARSSLPDSSFWAN 242
DB 321 AVDKY-----DVITFSRGPETADGRKPEVVAAGNMIIARAS---GTSWGP 366
QY 243 YNSKYAVMGTSNATPIVAGNVAQLREHFIKNGITP---KPSLIKALIAGATDVG-LG 298
DB 367 INDYTTAPGTSNATPVAAGIALLLQ---AHSWTPDKVKTALIEADIVKDEIADIA 423
QY 299 YPSGDQGMGRVTLDKSLNVAVNEATALTTGOKA-----TYSFQTOAGKPLKISLVWTD 353
DB 424 Y-----GAGRNVAKAIN--YDNVAKLVFTGYVANKSGTHQFVIGASFVATLWMDNA 476
QY 354 PGSTTASYTLVNDLDTLTAFAVNGQKYVNDPSYPDNNDGRNNVENVFINAPQSGTYTI 413
DB 477 N-----SDLDLYLDPNGNQV---DYSY-----TAYYGFEXVGYNPTDGTWTI 517
QY 414 EVQAVNPSGQRFSLAI 432
DB 518 KVVSY---SGSANYQVDV 533

RESULT 15
US-10-314-657-4
Sequence 4, Application US/10314657
Publication No. US20030175888A1
GENERAL INFORMATION:
APPLICANT: SHEN, Ben
APPLICANT: TANG, Gong-li
TITLE OF INVENTION: Discrete Acyltransferases Associated with Type I Polypeptide
TITLE OF INVENTION: Syntheses and Methods of Use
FILE REFERENCE: 054030-0021
CURRENT APPLICATION NUMBER: US/10/314,657
CURRENT FILING DATE: 2002-12-09
PRIOR APPLICATION NUMBER: PCT/US02/08937
PRIOR FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: US 60/278,935
PRIOR FILING DATE: 2001-03-26
NUMBER OF SEQ ID NOS: 214
SOFTWARE: PatentIn version 3.2
SEQ ID NO 4
LENGTH: 1237

TYPE: PRT
 ORGANISM: Streptomyces atroolivaceus
 US-10-314-657-4

Query Match 15.3% Score 346.5; DB 14; Length 1237;
 Best Local Similarity 30.4%; Pred. No. 3,1e-21;
 Matches 140; Conservative 55; Mismatches 154; Indels 111; Gaps 23;

QY	6	GIYKADVAQNN-----YGLYGGQYVAADTGLDTGRNDSMEAFRGKT--TAL	53
DB	213	GRVADLADSTAQIGACKWAEAGDTGQDVKVALDSDGADT-----EHPDLVGQVSDSAS	266
QY	54	YALGRTNNANDPNGHGTHTVAGSYLG-----NALNKGAPQANTVFOISIMSSGGLGPS	108
DB	267	FVPGEDDIA-DYNGHGHVASTITVGTGSASDGKERGVASGARLSTVGKVLNSEG--SGQES	323
QY	109	NLNTLFEQANNA--GARHTNSKGAFAVNCAYTAN--SRQVDEYVRNNDMTVLFAAGNE	162
DB	324	WI--IAGMEWAARDQKARLISMSLG---GGGDKNDPMSQAVDELSDTGALFVIAAGNG	377
QY	163	GPNSGTISAPGTAKNATITVGATENYRPSFGSIADNPNTIAQFSSRG-ATRDGRIRKPDVTA	221
DB	378	GPHS--ISSPGADSDALTVGA-----VDSTDTLADFSQSPRDGDGGLKPEITA	424
QY	222	PGTFILSARSSLAPDSFTFANVNSKYAVNGGISMATPIVAGNVAQLREHFINKRGITPPKP	281
DB	425	PGVDIVAAARSHYKRGSGY-----YTTMSGTSMATPHVAGVALLAAEHPDWTGTOLKE	477
QY	282	SLIKKALIAAGTVDGLYPSGDOGMGRVTLDRKSLNAVYNEATALTGQKATYSFQTQAG	341
DB	478	ALVSSAKATPA-----YTPYQAGAGRLDPAAPAAVHTTVFATTAYS-----GFHT---	521
QY	342	KPLKISLWTDAPGST--TASYTLVND---LDLVI--TAPNGQKYVGNDFSYPYDNMW	392
DB	522	-----WPKKPGETDVRTVYTNVGDAPVSLINLAVNGTVPAGL-----FSLSEDH--	565
QY	393	DGRNNEVNFINAPOSGTYTIEV-QAYNVPSGQRFSLAI	431
DB	566	-----YTVPAHGTAIVTTLTAALDKLADGQSVSAVI	595

Search completed: March 31, 2004, 16:33:59
 Job time : 34.1412 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:38 ; Search time 11.3246 Seconds

(without alignments)
3677.911 Million cell updates/sec

Title: US-09-985-689A-3

Sequence: 1 NDVARGIVADVAQNNGYL.....EVQAVNPDSGQRFSLAIWH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

PIR 78: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	502	22.2	1743	2 T18279	multidrug resistance
2	463.5	20.5	1905	2 T18267	intracellular alk
3	344	15.2	444	2 B83891	intracellular alk
4	315.5	13.9	442	2 A69587	microbial serine p
5	312	13.8	806	2 A41341	subtilisin-type pr
6	304.5	13.5	799	2 G83753	pyrolysins (EC 3.4
7	303	13.4	1398	2 T28159	serine proteinase
8	296.5	13.1	380	2 S11890	surface layer-asso
9	288	12.7	1345	2 T29090	alkaline serine pr
10	269.5	11.9	715	2 UC4908	high-alkaline seri
11	269	11.9	380	2 A49778	subtilisin (EC 3.4
12	265.5	11.7	420	1 S23407	alkaline proteins
13	265	11.7	534	1 JS0173	subtilisin (EC 3.4
14	258.5	11.4	419	1 S25835	subtilisin (EC 3.4
15	256.5	11.3	627	2 D75393	serine proteinase
16	256	11.3	757	2 C84120	subtilisin-type pr
17	255	11.3	401	2 I39974	serine proteinase
18	254.5	11.2	1331	2 A72647	probable surface
19	251	11.1	488	2 A19330	protease (import
20	248.5	11.0	1167	1 A35065	streptococcal C5a
21	246	10.9	894	2 F69730	cell wall-associat
22	243.5	10.8	378	2 A33973	high-alkaline seri
23	243	10.7	382	2 I39780	subtilisin (EC 3.4
24	243	10.7	379	1 A35742	subtilisin (EC 3.4
25	242	10.7	379	1 SUBSCL	subtilisin (EC 3.4
26	239.5	10.6	321	1 S27501	alkaline proteins
27	238	10.5	384	2 JC4802	alkaline proteins
28	236.5	10.5	402	1 JU0332	alkaline proteins
29	233.5	10.3	319	2 I39866	microbial serine p

30	229.5	10.1	519	2 S71451	halolysin R4 (EC 3
31	228	10.1	1052	2 T17093	intraluminal subtl
32	227.5	10.1	381	1 SUBBS	subtilisin (EC 3.4
33	227.5	10.1	381	2 JQ1487	subtilisin (EC 3.4
34	227.5	10.1	381	2 JH0778	subtilisin (EC 3.4
35	226.5	10.0	381	1 SUBSI	subtilisin (EC 3.4
36	226.5	10.0	905	2 F82734	serine proteinase
37	226.5	10.0	1433	1 A36734	serine proteinase
38	225	9.9	382	1 SUBBS	subtilisin (EC 3.4
39	224.5	9.9	272	2 A23624	subtilisin (EC 3.4
40	224.5	9.9	323	2 I39667	microbial serine p
41	224	9.9	274	1 SUBSD	subtilisin (EC 3.4
42	224	9.9	601	2 JC4576	serine proteinase
43	223.5	9.9	279	1 SUMTY	chitinase (EC 3.4
44	223	9.9	275	1 JCI085	subtilisin (EC 3.4
45	223	9.9	613	2 S75976	hypothetical prote

ALIGNMENTS

RESULT 1	
T18279	multidrug resistance, transport protein - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum	
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999	
C:Accession: T18279	
R:Shaulsky, G.; Loomis, W.F.	
Submitted to the EMBL Data Library, June 1996	
A:Reference number: Z18855	
A:Accession: T18279	
A:Status: Preliminary; translated from GB/EMBL/DBJ	
A:Molecule type: mRNA	
A:Residues: 1-1743 <SHA>	
A:Cross-references: EMBL:U60086; NID:g1399914; PID:g1399915; PIDN:AA03331.1	
A:Genetics:	
A:Gene: tagc	
Query Match	
Query Similarity 22.2%; Score 502; DB 2; Length 1743;	
Best Local Similarity 27.6%; Pred. No. 2,2e-24;	
Matches 162; Conservative 76; Mismatches 158; Indels 190; Gaps 24;	
QY	19 LYGGGVAVADPGLDPCR---NDS-----SMHEAFREKITALALAGTNNANPNNGH 68
DB	314 LRGGQGLSLADTGLDSHCFFSDSKYPIPLNSVNLNR-KVVTYTTSTSDSDKVDGH 372
QY	69 GTHVAGSVLG---NALN---KGMAPQANLVEQSIMDSGGIGLPSNINTLFSQAMNA 120
DB	373 GTHICGSAAGRPEDSVNISFSGLATDAKIAFPDLASGSSSLTP-PSDLKOLYPLXDA 431
QY	121 GARHTNSKGA---PVGATYANSRCQVDYV-RANDMTVLPAAGNEGPNSTIS--APG 173
DB	432 GARHCDMSGSVSEGYTGSISDPAIDPLTHPFIILRAAGN---NEOYLSTLTQS 488
QY	174 TAKNATVYGATE-----NYRSPFG-----192
DB	489 TAKNATVYGAGTHTHENYVLTDPVINYQSSVDINQSLICDPBSRYCNYTTAOCLESNA 548
QY	193 -----SIND-----NPNHIAQFSSRGATRDGRITPDVTAQPTFLISA 229
DB	549 TTGLASCCPTLLRKSVIDANTOPLLYNENNICSFSKSGPTHDGRMPALVARGEYITSA 608
QY	230 RSSLA-----PDSSPMANYSKAYVGTSMATPIVAGVVAOLREH-----FIK 273
DB	609 RSNGANTTDQCGDSL-PNNMALIA-ISTSMATSPAAATTLIKOYLUVGYPTGSIVE 666
QY	274 NRGITPKPSLIKALVIGA-----TDVGLGYPSGD-----QGMGRVTLTD 312
DB	667 SNKQPTGSLIKALMINNAQLNGTFLITSSITVPSNQVFNFAAGSLVQGMGAIKMS 726
QY	313 KSLNAVAVYNEAT-----ALTTGQKATYSF-----336
DB	727 NWLHVAVNNNSNNNKTSDDGITKFDGIGGLDLVAKPNQKESLSLGQNTSICFTYKPS 786

```

QY      33  --OTQAGKPELK---ISLWMTDAAGSTAYTYLVNDELIVTAPNCKQKVADPFYPADN- 350
Db      787  SSSNSGNNIIPRVATLVMTDPSPYAGAKFNLYNNLDTLMI-----YRDNQGITIFISNQ 841
QY      391  -----NMDCRRNVENVFVINAQSGSTYITIEQAVNNVSGPGRFS 428
Db      842  GSSSFLGLATPQDTLVNNEGIVANPEPMTYTFPMVAGTNVPGPQNF 889

```

RESULT 2

multiliding resistance protein - slime mold (Dictyostelium discoideum)
 C/Species: Dictyostelium discoideum
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T18267
 R/Shaulsky, G.; Kuspa, A.; Loomis, W.F.
 A/Submitted to: The EMBL Data Library, January 1995
 A/Description: An MDR transporter/serine protease gene is required for prestalk special
 A/Reference number: Z18850
 A/Accession: T18267
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1905 <SHA>
 C/Cross-references: EMBL:U20432; NID:g664839; PID:g664840; PIDN:AAA6212.1
 C/Genetics:
 I/Gene: tagB

Query Match	20.5%;	Score 463.5;	DB 2;	Length 1905;
Best Local Similarity	28.0%;	Pred. No. 7.8e-22;		
Matches 162;	Conservative 77;	Mismatches 156;	Indels 183;	Gaps 28

```

QY      19  LVGGGVVAADTGLTGR-----NDS-----SMEEARGGKTALYALGRNNAND--N 66
Dp      376  LRGGQLSLADTGGDGGSHCFPSDSKXPIPFNOVNEHNRKVITY-----TYHNEIYVN 430
QY      67  GHGHTHAGSVLGN-----ALN--KMAPQANLVQSIWDSGGIGGL-PSNLTFLPSQA 117
Dp      431  GHGHVCGSAAGRPEDSSWAISSFSGSLADAKIAFYDL--SSGSEPTPEPDYQOMKPL 488
QY      118  MNAGARIHTMSWGA---PVGAGYATNSHCQVEYV-RUNDMIVLPFAAGNEGPNSGTISAP 172
Dp      489  YDAGARHGSWSGSVSLQGYGGYSDSDAGGIDAPLYEYEPESILRAAGN-NELPASLIAQ 547
QY      173  GTAKNALTIVATE---NY-----RPSF-----191
Dp      548  ATAKNALTIVAEOTAHVNVYSDALEYYDESDNANFORPCLPDKXYCNNTAKCSESVNV 607
QY      192  -----GSIAQ-----NPHIAQFSRRGATRDGRKIPDVTAFTFILSAR 230
Dp      608  KGALQCCPASIKQNASDSFTTQPOFQENNMNGSFSKGPETHGRKIPDIVAPGSYITSAR 666
QY      231  SS-----LAPDSSFMANNSKRAYMGISMATPIVAGNVACLREHF-----IKX 274
Dp      668  SNGENSTDOCGDSSL-PNANGLWS-LSGISMATPATAATTLRQYIVDGYFPTGESVEE 725
QY      275  RGIITPKESLIKALIIAGATDVGLGY-----PSGD-----QGWRTYTLKSL 315
Dp      726  NKLLPTSLIKALMINNAQLNLTGYFWSASSTNPENALFEQJINGANLIIQMGALRRNNWL 785
QY      316  NVAYVN-----EAT---ALTIGQAKTYSFCQJ-----AGK 342
Dp      786  YVSSSNFTPPSRMIGIGGLGKQKAKLEWKEDSLISGLNKSXYCFYKPPSSSSGSGGGGGT 845
QY      343  P-LKISLVMTDAGSGTASYATVINDDL-----VITAPN--GQKYVNDSEFYDN 390
Dp      846  PRIVATLVMTDPPSYGAKFNILVNNLDLLLSNDDDSITITENSGSLQPRAGKVAQF--- 902
QY      391  NMDGRNNVEVFTNAPOSGTYTTEVQAVNVBPGQPS 428
Dp      903  --DTLNVVEGIINFTKAMNYKFTTIGATVNPVGPOKFS 938

```

RESULT 3

Intercellular alkaline serine proteinase apx [imported] - *Bacillus halodurans* (strain C) SpecDes: Bacillus halodurans
C Date: 01-Dec-2000 #sequence: revision 01-Dec-2000 #text: change 15-Jun-2001
C Accession: B83891
R Takami, H., Nakasegawa, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fujii, F., Hir
Nucleic Acids Res. 28, 4317-4331, 2000
R Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
Reference number: A83650, WUID: z0515582, PMID:11056132

Query Match	15.2%;	Score 344;	DB 2;	Length 444;
Best Local Similarity	33.6%;	Pred. No. 6.1e-15;		
Matches 110;	Conservative 47;	Mismatches 110;	Indels 60;	Gaps 17;

```

QY      11 DVACNNNGILGQGGGVVAVADTGLTGTGRNDSMHEAFKGLITALY-AGRTNADPNHG 69
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      135 EVIENGBSTLTKDVTIAVIDTGI-----YPHEDLBGRIRAFVDFVNGREPYDDNGH 187

QY      70 THVAGSTVLGNALN-----KGMAPQANLVFGSIMDSGGGLGPENNTATLFGQAM----- 118
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      188 THCGDAAGAGSASDQYRGAPAPAEANIVGVKLKQ-QMGSLESIMQGV--EMCIQYNE 243

QY      119 -NAGARIHTMSWGAFFVNG-AYTANSRQVDEXVR-----NNDMTVLFPAGNEGPNSTIS 170
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      244 EHPDPPHIITMS--LGGQALPYENEGQDPVVRILVEEMANAGITVCVAAGSGPDAQITIA 301

QY      171 AFGAKAALITVGALENTRPFSGSADNPNHIAQSSSGATRDGRLKPDVTAPGTFILSAR 230
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      302 SPGVSEVITVGAALDD-RDITDREDD--VAPSSRGFTIYGKPKDILAPGVNATSLR 357

QY      231 SSLAPDSSF-----WANYNSKYAMGGSNATPIVAGNVAAQLBHRFIKNRGITPEPSLIK 285
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      358 S---PNFSDYKIKIGSRVSGSHYTMGSGSNATPVCAGVALLMLQH---EPYLTDE--VK 409

QY      286 AALTAGATD-----VGLGYPSGD 303
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      410 TRLMEG--TDRMADRDENVYGAGYTSAR 435

```

RESULT 4

intracellular alkaline serine proteinase aprx - *Bacillus subtilis*
C/Species: *Bacillus subtilis*
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C/Accession: A69367
R/Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertele, C.; Bron, S.; Broillett, S.; Bruschi, C.V.; Caldwell, B.; Capriano, V.; Carter, N.M.; Ch A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallie, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsteppel, S.; Hosono, S.; Hullo, M.F.; Iech, J.; Konigsgesin, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, Koetter, P.; Konigsgesin, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A/Authors: Lander, U.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maues, Y. M.; Ogawa, K.; Ogatawara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A/Authors: Schlecht, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, T.; Sekowska, A.; Sekowski, akuchil, M.; Tamakoshi, A.; Tanaka, T.; Terpeita, P.; Tognoni, A.; Tose, V.; Uchiyama, T.; Winiers, P.; Wipit, A.; Znamenski, H.; Yamane, K.; Yasumoto, K.; Yate, K.; Yoshida, A/Authors: Yoshikawa, H.F.; Znamenski, H.; Yoshikawa, K.; Yano, H.; Dancib, A.
A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A/Reference number: A69580; MUID:98044033; PMID:9354377
A/Accession: A69367
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: G83753
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-799 <STO>
A;Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA04550.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: vpr
C;Superfamily: microbial serine proteinase vpr; subtilisin homology
C;Keywords: hydrolase; serine proteinase
F;1-29/Domain: signal sequence #status predicted <SIG>

Query Match 13.5%; Score 304.5; DB 2; Length 799;
Best Local Similarity 23.9%; Pred. No. 4.9e-12;
Matches 149; Conservative 59; Mismatches 152; Indels 263; Gaps 28;

QY 18 GLVGGGVAVADTGLDTGRNDSMEAFRGKITALVALGR-----TNN-----AN 63
DB 171 GYTGEGITVALDITGVDTYHPD-----LVHAFGDYKQWDFINNDPQETPRG 218
QY 64 DPNQ---HGTHVAGSYLGNALKKMAPOANLVFQSIIMSSGGLGSPNLT--FSQA 117
DB 219 DPGIETTHGTHVAGSYLGNALKKMAPOANLVFQSIIMSSGGLGSPNLT--FSQA 274
QY 118 WNAGARHTMSWGPVAVGATYANRQVDEYVRNNDMTVLPAAENEPNSGTSIAPSTAK 177
DB 275 VQDADIMNLSLGTNLNDPFAISILD--WAMEGVAVTNSNGSPNMTVSGPSTSD 333
QY 178 AITYGAT-----ADNP-----ENYRPSFGST----- 194
DB 334 AIVGATRLPYNKYKASVFTSDGIDPSADIMGFSPDELLDGETYEYAFAGLQKPGD 393
QY 195 -----ADNP-----ADNP----- 198
DB 394 FEGDVAGKIALIVRGIIPFEYKAKENAKAGAVGAIYNNVAGVCPVPGALPTIMLSN 453
QY 199 -----NHIAQPSRG--ATRDGRIKPDVATPGLTILSA 229
DB 454 EDGLKNELENGONTVFSIEPKLVGETVADFSSGPMHTMKKPDVASGVAIVST 513
QY 230 RSSLAPSSFPANYSKAVWGTSNATPIVAGVAGQREHFIKNGITPKSLIKALIL 289
DB 514 IPTHQPDPT--GYGSRQ-----GTSNASSPVAAGALLLEAH--PNMGV---DHVRAALM 562
QY 290 AGATDV---GLGYSGDQGWGRVTLDSLVNAVYNEATALTG----- 329
DB 563 NTAENLVDENGNRFPNHTQAGSIRI-----VDALIESETLVTPGSHSFGFTKRGQVE 617
QY 330 -----OKATYSFGTQ--AGKP-----LKISLVWTDAPGSTSYTLVNDLVIRA 373
DB 618 RQHTTINLSNKRKITYDFDVQFAGNPDGIKYSKILRVPGKT----- 661
QY 374 PNGQKYYGN-----DFSYPYDNN--WDGNNVY--NVEINA---PQSGTYITIEV- 415
DB 662 ---QKINFNQVDARKIDPGR--YEGTIVSDGSQIVETPILRVSESDYRVTTFDLDID 717
QY 416 -----QAYNVSPGPRFSLAI 431
DB 718 ENGVLFGSAY--LPGAGEEFGIWI 739

RESULT 7
T28159
Pyrolysin (EC 3.4.-.-) - *Pyrococcus furiosus*
C;Species: *Pyrococcus furiosus*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 24-Oct-2000
C;Accession: T28159
R;Voorhorst, W.G.B.; Eggen, R.T.L.; Geerling, A.C.M.; Platteuw, C.; Sizeren, R.J.; Vos,
J. Biol. Chem. 271, 20426-20431, 1996

A;Title: Isolation and characterization of the hyperthermostable serine protease, pyrolysin
A;Reference number: Z20481; MUID:56355370; PMID:8702780
A;Accession: T28159
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1398 <VOO>
A;Cross-references: EMBL:U55835; NID:g1556462; PID:g1556463; PIDN:AAB09761.1
A;Experimental source: DSM3638
C;Genetics:
A;Gene: pls
C;Keywords: hydrolase; serine proteinase

Query Match 13.4%; Score 303; DB 2; Length 1398;
Best Local Similarity 28.1%; Pred. No. 1.3e-11;
Matches 128; Conservative 37; Mismatches 154; Indels 136; Gaps 15;

QY 21 GQGVAVVADTGLDTGRNDS-----SMHEFRGKITALVALGRTNANDPN-- 66
DB 301 GNGIDIAVDITDIDTDEVPGLQVNVTVDAVFSTYIGPLN--YLAET---DPNGE 354
QY 67 -----GHGTHVAGSYLGNALN-----KG 84
DB 355 YAVFGMDGHHGTHVAGVAGYDSNDAMDMLSVSGEMVEFRLYQMDYTNVTTDTVQ 414
QY 85 MAPQANLVFQSIIMSSGGLGSPNLTLSQANAGARHTMSNG--APVNGATYANR 142
DB 415 VAFGAQMAIRVLSDG--RGSWMDIEGMYATRHADVISMSLGNAPYLDGTPESV 472
QY 143 QVDEYVRNNDMTVLPAAENEPNSGTSIAPSTAKNATVAGT----- 185
DB 473 AVDELTEKYGAVFIAAGNEPGLNIVGSPGATKAITGAAVPIVNGYVQALGPD 532
QY 186 ---NYRPSFGSINDPNHIAQPSRGATRDGRIKPDVTPGFIISARSLAPDSFMA 241
DB 533 YVGFYPPAYTNV-----RIAFSSRGPRIDGELKPVVAPGVGIYSSLPMWIGADF-- 585
QY 242 NYNSKYAAMGTSNATPIVAGVAGQREHFIKNGITPKSLIKALILAGT-----DV 295
DB 586 -----MSGTSNATHTVSGVALLISG--PKPEGIYNNDDIIKVLTSAGATWEGDPT 636
QY 296 GLGYPSGDQGWGRVTLDSLVNAVYNEATALTGQKATYSFGTQAGKPLKISLVWTDAPG 355
DB 637 GQKTELDQGHGLNVVTKWEI-----LKAINGTTLPIVDMADKSY 678
QY 356 STTASYTLVNDLVDLVITAPNG-----QKYVGN 382
DB 679 SDFAEYLGAV-DVIRGLIARNSIPDIVEMHKKYVGD 712

RESULT 8
S11890
serine proteinase (EC 3.4.21.-) precursor, extracellular - *Xanthomonas campestris* pv. c
N;Alternate names: subtilisin-related proteinase
C;Species: *Xanthomonas campestris* pv. *campestris*
C;Date: 21-Nov-1993 #sequence_revision 07-Feb-1997 #text_change 03-Dec-1999
C;Accession: S11890
R;Liliu, Y.N.; Tang, J.L.; Clarke, B.R.; Dow, J.M.; Daniels, M.J.
Mol. Gen. Genet. 220, 433-440, 1990
A;Title: A multipurpose broad host range cloning vector and its use to characterise an
A;Reference number: S11890; MUID:90251253; PMID:2187155
A;Accession: S11890
A;Molecule type: DNA
A;Residues: 1-580 <LIU>
A;Cross-references: EMBL:X51635; NID:g48533; PIDN:CAA35962.1; PID:g48534
A;Experimental source: *Xanthomonas campestris* pv. *campestris*
A;Note: the sequence from Fig. 4 is inconsistent with that from Fig. 3 in having 205-A1.
C;Superfamily: subtilisin; subtilisin homology
C;Keywords: extracellular protein; hydrolase; serine proteinase
F;1-33/Domain: signal sequence #status predicted <SIG>
F;168-423/Domain: subtilisin homology <SHT>

Query Match 13.1%; Score 296.5; DB 2; Length 580;
Best Local Similarity 29.0%; Pred. No. 1e-11;

Matches 137; Conservative 52; Mismatches 148; Indels 135; Gaps 26;

QY 21 GGGGVAAVADTGL-----DTGRNDSMEAFRGKITALYALGRTNAND-----64

DB 168 GSGTVAIVADTITSHADINANIILAGYDFISDATTARONGSDSNAADGDIYANNECGA 227

QY 65 -----PNEGTHVAGS-----VIGNALINKGMAPQANL-----VFQSIMDS 99

DB 228 GTPAASSMHGTHVACTVAATNNTTGVAGTAVGAKVYVRVLGCGGSLIDIAIATYMA 287

QY 100 SGG-IGGLPSNLN--TLFSQAMNAGARIHTNSMGAPVNGAYTANSRQVDEVYRNNDMTVL 156

DB 288 SGGTVAIPANAPAVINMSLGGGSCSTTQNA--INGAVSRGT-----TVV 334

QY 157 PAAGNCGPN--SGTISAPGAKAIVTGATENRPSFG--ADNPNNHIAQSSSGATRDRI 215

DB 335 VAAANDASVSG--SLPANCANVIAVAT-----TSAGA-----KASISNFGT-----375

QY 216 KPDVTAPEGTFILSARSS--LAPDSFWANYNSKYAYMGSTMATPVAQNVAAQLREHPLK 273

DB 376 GIDVSAFGSILSTLNSGTTTSGASASYN-----GTSMASSHVAGVALVQS--VA 426

QY 274 NRGITK--PSLIK--AALIAAGTDVGLG-----PSGQGGRTVLDKS 314

DB 427 FYALTPAVETTLKNTARALPGACSGCGAGIVNADAATAALINGSGGGGGGNTLNG 486

QY 315 LNVAYVNEATALTGOKATYFOTOAKPLKSLVWTDAPGSTTASVTLVNDLIVT---371

DB 487 TPTVGLGA---TGAEINATITVPAQSG--TLVTYTSGGG-----GADLYVRAG 531

QY 372 TAANGKQYVNDSTPYDNNWDRNNVENVFINAPSGTYTTEVQAYNPNPSG 423

DB 532 SAPTDSAYT---CRPYRS-----GNAETCTITAP--SGTYVYRLKAYSTPSG 573

RESULT 9

T29090

surface layer-associated STABLE proteinase - Staphylothermus marinus

N:Alternate names: hyperthermostable proteinase

C:Species: Staphylothermus marinus

C>Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #ext_change 02-Sep-2000

C:Accession: T29090

R:Mayr, J.; Lipas, A.; Kellermann, J.; Eckerskorn, C.; Baumeister, W.; Peters, J.

Curr. Biol. 6, 739-749, 1996

A:Title: A hyperthermostable protease of the subtilisin family bound to the surface layer

A:Reference number: Z20539; MUID:96385442; PMID:8793300

A:Accession: T29090

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1345 <MAY>

A:Cross-references: EMBL:U57968; NID:g1374755; PID:g1374756; PIDN:AA802323.1

C:Experimental source: strain F1

C:Function:

A:Description: probably serves an exodigestive function related to the organism's energy

A:Note: stoichiometric S-layer component

Query Match 12.7%; Score 288; DB 2; Length 1345;

Best Local Similarity 25.5%; Pred. No. 1.2e-10;

Matches 120; Conservative 66; Mismatches 152; Indels 130; Gaps 20;

QY 46 FRGKITLVALGRTNANDPNEGTHVA-----GSVIGNALN-----KGMAPQANLV 92

DB 445 YQGRVYAL-----VSDFHGHTSVATVIAISRGVLYDLYGDKLYRIMGVAPAKI- 495

QY 93 FQSIMDSGGGLGGLPSNLTLFSQAMNAG-----ARITNSW 129

DB 496 -----AGSDAWLLGNI--LVLEAWLAGENITVEEDGVYVSLDPGPHRADITISNSH 546

QY 130 GA-----PVGAYTANSROYDEVYRNNDMTVLPAAGNCGPNSTISAPGTAK 176

DB 547 GSIYINFWLQGFPGIDYRSSFMDEILAIIRYLLIGDHVTIVFAAGNEGPYSGYSCAGAGTGL 606

QY 177 NAIIVGATE--NYPSPFGSIADNPNNHIAQSSSGRATDGRIGPDVTAPEGTFILSARSSLA 234

DB 607 LVITAGASTIMDTYRIVYEGYAGVADIVIPSSSGPGCGQPKRDYVNIAGF-----657

QY 235 PDSSFWAN-----YNSKYAYMGSTMATPVAQNVAAQLREHPLKNGRITPKSLIRA 286

DB 658 -----EMASTRTIDRGYGAQPDVFGGTSEATPTSSGLLVPOAYKEVNTTDDPYTAKI 713

QY 287 ALIAGATDVLGYPSGQGWGRTVLDKSLNVAVNEATA--LTGOKATY--SFQOAGKP 343

DB 714 ILKSAADI--WTPAPSGSGRDALKADDTVISEIEMIAVSGIOGAEPLFNYYTDPGPY 771

QY 344 LKISLVN--TD-----APGSTASYTL--VNDLIVTAPNG--QKYVNDSTPYDN 390

DB 772 IGYVLPYADTDYGVVYKPGS--SKNFTLVNGAVLSAMNTVLYKEYTVVDGVVDYSG 830

QY 391 -----NMDGRNVE-----NVRINAPSGTYTTEVOAYN 419

DB 831 LFLKVPKXVSGADYVEVVYQLENNTPPGVFLKTPIDPLHALIISAYD 880

RESULT 10

JC4908

alkaline serine proteinase (EC 3.4.-.-) I precursor - Alteromonas sp.

N:Alternate names: subtilase

C:Species: Alteromonas sp.

C>Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #ext_change 08-Oct-1999

C:Accession: JC4908

R:Tsujiho, H.; Miyamoto, K.; Tanaka, K.; Kaidzu, Y.; Imada, C.; Okami, Y.; Inamori, Y.

Biosci. Biotechnol. Biochem. 60, 1284-1288, 1996

A:Title: Cloning and sequence analysis of a protease-encoding gene from the marine bacterium *Alteromonas*

A:Reference number: JC4908; MUID:97141200; PMID:8987544

A:Accession: JC4908

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-715 <TSU>

A:Cross-references: DBJ:D36600; NID:g1536787; PIDN:BA18912.1; PID:d1019647; PID:g2160

A:Experimental source: strain O-7

C:Comment: This enzyme belongs to class I subtilisin-like family. It is a chelator-sens

C:Genetics:

A:Gene: aprI

C:Superfamily: subtilisin homology

C:Keywords: hydrolase

F:1-40/Domain: signal sequence #status predicted <SIG>

F:41-150/Domain: amino-terminal propeptide #status predicted <AMP>

F:151-456/Product: alkaline serine protease I #status predicted <MAT>

F:182-452/Domain: subtilisin homology <SST>

F:497-715/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F:239-294,335-372,478-481/Disulfide bonds: #status predicted

Query Match 11.9%; Score 269.5; DB 2; Length 715;

Best Local Similarity 25.6%; Pred. No. 7.8e-10;

Matches 125; Conservative 44; Mismatches 164; Indels 155; Gaps 20;

QY 21 GGGGVAAVADTGL-----LDGRNDSMEAFRGKITALYALGRTNANDP-----65

DB 182 GGGVVAAVADTGYRPHLDLNNILPGYDMISNTFVANDGARGDADDPGDAYTRGEGCT 241

QY 66 -----NGGTHVAGSVLGNALN-----KGMAPQANLVFOSIMSGGLGLP 107

DB 242 DSSGQPVPRADQDSKGTHTVAGTVAATNNGEVAAGVADAKVYPRVL--GKCGGLT 298

QY 108 SNLNTLFSQAMNAGARIHTNSWGA PV-----NGAYTANSROYDEVYRNNDMTVLPAAG 160

DB 299 SDIADGIIWASGSDRPANAPVAVINMSLGGGAGSATTQNAINQARNNGTIVIAAG 358

QY 161 NEGPNSGTISAPGAKAIVTGATENRPSFG--ADNPNNHIAQSSSGATRDGR-----214

DB 359 NDNDNSA-----NTPPG-----NNGVYNAVSG--RDGSRATYSN 392

QY 215 --IKPDVTAPEGTFILSARSLAP-----SSFANANTSKYAYMGSTMATPPIYAGN 263

DB 393 YGANIDVAPG-----GAGSPADDPGILSTTHNSGAGPSNDSTHYEGGTSMAPAHVAGV 447

QY	264	VAQREHETIKRGGTTPKPSLIKALLIAGATVGLGPEBGDGMGRVTLDSLNAAVNEA	323
Db	448	AAIIKQ-----AKPSATPDEVEITLIKNTTSPFAGSCSCGTVADAA--AAVNEA	495
QY	324	-----TALT--TGOKATYSEFOTQAGKPEIKISLIVTDAFSGTASYT	362
Db	496	LGDVVTPTPTGNTLBDGVAKTGLSGAAGSNQFFTPDVPAGK-----TNVTFT	541
QY	363	L--VNDDLVITAPNGQKYVGNDFSYEDNNMDGR-----NNVENVFIAAQSGTYTTEV	415
Db	542	MSGGTGDDADLYVK-----LG---SPTSSSYDORPYEGGNAEYCSFDPDAQATYHMI	591
QY	416	QAYVTPSG	423
Db	592	NGRYAYSG	599

RESULT 11

high-alkaline serine proteinase (EC 3.4.21.2) - precursor - *Bacillus alcalophilus* (strain N) | Alternate names: subtilisin homolog, high-alkaline C | Species: *Bacillus alcalophilus* C | Date: 12-Mar-1994 | #sequence_revision 24-Feb-1995 | #text_change 20-Jun-2000 C | Accession: A49778, JCI1244 R | van der Laan, J.C.; Gerritse, G.; Mulleners, L.J.S.M.; van der Hoek, R.A.C.; Quax, W.J. Appl. Environ. Microbiol. 57, 901-909, 1991 A | Title: Cloning, characterization, and multiple chromosomal integration of a *Bacillus* A | Reference number: A49778, PMID:91282483, PMID:2059048 A | Accession: A49778 A | Molecule type: DNA A | Residues: 1-380 <VAN> A | Cross-references: GB:ME5086; NID:g142456; PIDN:AAA2212.1; PID:g142457 A | Experimental source: strain P392, ATCC 31408 A | Note: amino end of mature protein confirmed by peptide sequencing R | Itakami, H.; Kobayashi, T.; Kobayashi, M.; Yamamoto, M.; Nakamura, S.; Aono, R.; Horikawa, T. Biotechnol. Biochem. 56, 1455-1460, 1992 A | Title: Molecular cloning, nucleotide sequence, and expression of the structural gene A | Reference number: JCI1244; PMID:93043753; PMID:1368952 A | Accession: JCI1244 A | Molecule type: DNA A | Residues: 1-195, 'S', 197-380 <TAK> A | Cross-references: GB:D1315; NID:g216231; PIDN:BAAN02442.1; PID:g216232 A | Experimental source: *Bacillus alcalophilus* Veddler, ATCC 21522 (*Bacillus* sp. 221) C | Superfamily: subtilisin, subtilisin homolog C | Keywords: hydrolase; serine proteinase; zymogen F | 1-27/Domain: signal sequence #status predicted <SIG> F | 28-111/Domain: activation peptide #status predicted <PRO> F | 112-380/Product: alkaline serine proteinase #status predicted <MAT> F | 134-340/Domain: subtilisin homolog <SBT> F | 143,173,326/Active site: Asp, His, Ser #status predicted

```

Query Match 11.9% Score 269 DB 2 Length 380;
Best Local Similarity 31.8% Pred No 3.6e-10;
Matches 97; Conservative 35; Mismatches 99; Indels 74; Gaps 14;

QY 13 AONNYGLYGQGVAVADTGLDTGRNDSMEAEFGKLTALYALGRTNANDPNHGCHTV 72
Db 126 AAHNGLTGGGVKAVLDTGIST-----HDLNIRGASGVPEP-STDGNHGCHTV 177
QY 73 AGSV--LGNALN-KMARPQANLVPOSIMDSGGGLGSLNLTLPQAMNAGARIHTSM 129
Db 178 AGTIALNNSIGVLGVAPVLAELVAKVLGASG--SGSVSIIQGLEWAGNNGMRYANLSL 235
QY 130 GAPVAGATANGROYDEVYRNNDMTVLTPAGGEGNSGRTISPTAKAALITGATENRP 189
Db 236 GSP---$PSATLEQAVNKAISRGVTLVAASNSG--AGSISIPALYANMAVGA----- 285
QY 190 SFGSLADNPNHTAQSSRCATRDGRIKPDVTNAPGFFIILASRSLAPDSFWANYSKAY 249
Db 286 -----DONNNAASFSQYAGL-----DIAPGVNVGSTYP-----GSTAS 321
QY 250 MGGTSMATPIVAGNVA-----QLRHFINKRGITPEPSLKAALIAAGATDVL 297

```

Db	322	LNQTSNATHTHVAGAAALVKQKNPSMSNVQIRNH-LKN-----TATSLGSLNL--	367
QY	298	GYPSG 302	
Db	368	-YGSQ 371	

RESULT 12

253397
 subunit11sin (EC 3.4.21.62) 1 precursor - *Bacillus* sp. (strain TA39)
 C.Species: *Bacillus* sp.
 C.Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 18-Jun-1999
 C.Accession: S23407
 R.Narinx, E.; Davail, S.; Feller, G.; Gerday, C.
 Biochim. Biophys. Acta 1131, 11-113, 1992
 A>Title: Nucleotide and derived amino acid sequence of the subunit11sin from the antarctic
 A.Reference number: S23407; MUID:92256481; PMID:1581352
 A.Accession: S23407
 A.Molecule type: DNA
 A.Residues: 1-420 <NAB>
 A.Cross-references: EMBL:X62369; NID:G40200; PIDN:CAA44227.1; PID:G40201
 C.Genetics:
 A.Gene: sub1
 C.Superfamily: subunit11sin, subunit11sin homology
 C.Keywords: extracellular protein; hydrolase; serine proteinase
 F.1-33/Domain: signal sequence #status predicted <SIG>
 F.24-111/Domain: propeptide #status predicted <PRO>
 F.112-420/Product: microbial serine proteinase #status predicted <MAT>
 F.135-374/Domain: subunit11sin homology <SBr>
 F.145,185,350/Active site: Asp, His, Ser #status predicted

Query Match 11.7%; Score 265.5; DB 1; Length 420;

Matches 92; Conservative 46; Mismatches 120; Indels 59; Gaps 15;

Qy 21 GGGVVAVADTGLDGRNDSMHEAFRCKITAL--YALGRT--NNANDPNHGTHVAGS 75

Db 136 GGGINIAVLDTGVTN-----HPDLRNVEQCKDFTVGTTYTNNSCTDRQGHGTHVAGS 189

QY 76 VL-----GNALNKGMAPQANLVFQSIM--DSSGGLGGLPSNLTLEFSQAWNAGARIHTN- 127

Db 190 ALADGGTNGV - YGVAPDADLWYKVLGDDGSGYADDIAAARIHAGDQATALTNTKVINM 248

QY 128 SWGAPVNGAYTANSRQVDEYVRNDMTVLFAAGNEGPNSGTISAPGTAKNAITVGATENY 187

Db 249 SLGSSGESSLITNA--VNYSYNGVLIIAAGNSGPYQGSIGYPGALVNAVAVALEN- 304

188 RPSFGSIADNPNHIAQFSSRG-----ATRDGRIKPDVTA¹GTFILSARSSLPDSSF 233

Db 305 -----KVENGTYRVADFSSRGYSWTDGDYAIQKGV--EISAFGAIIYST----- 347

240 WANNYSKYAYMGTSMATPIVAGNVAQLREHFIKNRGITPKPSLIKALITAGATDVGLGY 299

348 W--PDGGYATISGTSMASPHAGLAAKIWAQYPSASNVDVRGELYRAY--ENDILSGY 402

300 PSG-----DOGWGRVTL 311

Db 403 YAGYGDDEASGFGPATV 419

RESULT 13
TS0173

N: Alternate names: alkaline serine exoproteinase A
alkaline protease (EC 3.4.21.-) A precursor - *Vibrio alginolyticus*

C:\species: *Vibrio alginolyticus*
C:\Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999

C:/Access10n: JSU173
E T : Bobb S M : Woods D R
F:Deane S M : Bobb

Gene 76, 281-288, 1989

A;Reference number: JS0173; MULD:89326126; PMLD:2546861
 A;Accession: JS0173

A;Molecule type: DNA

A:Residues: 1-534 <DEA>
 A:Cross-references: GB:M25499; NID:G155250; PIDN:AA27550.1; PID:G155251
 C:Comment: This protein is a calcium-dependent and sodium dodecyl sulfate-resistant pro
 C:Superfamily: subtilisin; subtilisin homology
 C:Keywords: calcium; hydrolase; serine protease
 F:1-21/Domains: signal sequence #status predicted <SIG>
 F:122-534/Product: alkaline serine proteinase A #status predicted <MAT>
 F:171-378/Domains: subtilisin homology <SBT>
 F:271-273,297-299/Region: S1 specificity crevice #status predicted
 F:180,213,363/Active site: Asp, His, Ser #status predicted

Query Match 11.7%; Score 265; DB 1; Length 534;
 Best Local Similarity 26.3%; Pred. No. 1e-09;

Matches 123; Conservative 65; Mismatches 152; Indels 128; Gaps 24;

```

QY 7 IYKADVAQ-----NNY--GLYGGQVAVAVADGLDGRDSSMHEAF 46
DB 137 IYKADVAQ-----NNY--GLYGGQVAVAVADGLDGRDSSMHEAF 190
QY 47 RGRITALLY-ALGRTNNANDPNHGHTVAGSVLGNALNKGAPQANLVFGSIM--DSGGL 103
DB 191 GGRSVGVDFVDDADADPCNGHGTVAHTIGSLY--GVAKNVLVGVRLVSCSGSGST 248
QY 104 GGLPSNLNTLFSQANAGARIHTNSGAPVNGAVTANSRQVDEYRNNDMTVLPAAGNG 163
DB 249 SGVIAGDVWVAANA--SGPSVANMSLG--GGQSVLADSAVQSAVQSG-VSFMLAAGNSN 302
QY 164 PMSGTISAPGTACATVGTATENYRPSFGSIADPNHIAQFSSRGATRDGRITKPYTAPG 223
DB 303 ADACNYS-PARVATGVTVGST-----TSTDARSSFSWMSGC-----VDVFAPG 344
QY 224 TFLLSARSSLAPDSSEFWANYSKYATMGTSNATPIYAAGVAQLRHFIRKNGITPKPRL 283
DB 345 SQIKSA-----W--VDGGYKXTISGTSNATPVHAG-VVAL--VLQENSSVS--PSQ 387
QY 284 IKALILAGA-----TDVGLGYVSG--DQGMGVTLDKSLINVA 319
DB 388 VEHILVERASTGKVTDRGSVNLVSLIDADCGCGGDPDPPEGLTSGVPVS--- 444
QY 320 VNEATALTITGQKATYSFQIQAGKPLKISLVMTDAPSGTASTAYLVNDLVLITAPNGQY 379
DB 445 ---GLSGSSQVAVYVYDVAGQRLTYQM-----YQGSADALYLRF--GAK- 486
QY 380 VGNDFSPYDNMWDGR---NNVENVFINAPOSCTYIEVOANYVPSG 423
DB 487 -----PTINAMDCHRFKYNNECTVSKATQSGRIHWMIQGISNMSG 527

```

RESULT 14

S25835
 Subtilisin (EC 3.4.21.62) precursor - *Bacillus* sp. (strain TA41)

C:Species: *Bacillus* sp.
 C>Date: 22-Nov-1993 #sequence_revision 20-Feb-1995 #text_change 18-Jun-1999

C/Accession: S25835
 R:Davidall, S.; Feller, G.; Narinx, E.; Gerday, C.

Gene 119, 143-144, 1992
 A:Title: Sequence of the subtilisin-encoding gene from an antarctic psychrotroph *Bacillus*

A:Accession: S25835; MUID:93012966; PMID:1398082

A:Molecule type: DNA
 A:Residues: 1-419 <DAV>

A:Cross-references: EMBL:X65533; NID:G40198; PIDN:CAA5096.1; PID:G40199

C:Superfamily: subtilisin; subtilisin homology
 C:Keywords: extracellular protein; hydrolase; serine protease

F:1-23/Domains: signal sequence #status predicted <SIG>
 F:24-110/Domains: propeptide #status predicted <PRO>

F:111-419/Product: microbial serine proteinase #status predicted <MAT>
 F:145-373/Domains: subtilisin homology <SBT>

F:144,184,359/Active site: Asp, His, Ser #status predicted

Query Match 11.4%; Score 258.5; DB 1; Length 419;
 Best Local Similarity 31.1%; Pred. No. 2e-09;
 Matches 82; Conservative 40; Mismatches 99; Indels 43; Gaps 12;

```

QY 21 GGGVAVAVADGLDGRDSSMHEAFRGRITALLYALGRT---NNANDPNHGHTVAGSV 76
DB 135 GAGINIAVADLTGVYTHNHPDLSSNVEQCKD-----FVGNFTDSCCTDQGHGHTVAGSA 189
QY 77 LGNANL---KGAPOANLVFGSIM--DSGGIGLPSNLNTLFSQANAGARIHTN--SW 129
DB 190 LANGTGSGVYGVAPKADLWAVYVLDDGGVADDAEALRHAGDQATALTNTKVTINMSL 249
QY 130 GAPVNGAVTANSRQVDEYRNNDMTVLPAAGNGPNSGTISAPGTACNATVGTATENYRP 189
DB 250 GSGSESLINNA--VD-YADKGVLLIAAAGNSGPKRPGTGYGALVNAVVAALNN--- 303
QY 190 SFGSIADPNHIAQFSSRGATRDG-----RIKPYTAPGTFLLSARSSLAPDSFWAN 243
DB 304 ---TIQGTYRVADFSSRGHRTAGDYVLQKGDVEISABGAAYST-----W--F 348
QY 244 NSKYAVMGTSNATPIYAAGVAQL 267
DB 349 DGGYATISGTSNATPVHAG-VVAL- 372

```

RESULT 15

D75393
 serine protease, subtilase family - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C/Accession: D75393
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, H.O.; Vamathavan, J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Mc

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: D75393
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-627 <MHI>

A:Cross-references: GB:AE001990; GB:AE000513; NID:G6459214; PIDN:AA11026.1; PID:G64592;

A:Experimental source: strain R1
 A:Gene: DR1459
 A:Map position: 1

Query Match 11.3%; Score 256.5; DB 2; Length 627;
 Best Local Similarity 24.8%; Pred. No. 4.5e-09;

Matches 104; Conservative 61; Mismatches 161; Indels 93; Gaps 17;

```

QY 18 GLYGGQVAVAVADGLDGRDSSMHEAFRGRITALLYALGR---TNANDPNHGHTV 72
DB 161 GFKGQNTIRIHLDDGIDPIS---HPELAGKVAAFQEFNGEGDRVSSQPHDTTGHHT 214
QY 73 AGSVLGNALNKGAPQANLVFGSIMDSSG---LGLPSNLNTLFSQANAGARIHTN 127
DB 215 AGLVGSRY--GVAPKAVISALVLPNNEGFAQVIAQMGVYLDPDNNADTDQADAVNM 272
QY 128 SWGAPVNGAVTANSRQVDEYRNNDMTVLPAAGNGSGSTISAPGTACNATVGTATENY 187
DB 273 SLGIRP--GTWNEFVVPVNNMLKAGVVPV-FALGNFGPAAGSGSPGMLPCALIGVGA--- 325
QY 188 RPSFGSIADPNHIAQFSSRGATRDGRI---KPYTAPGTFLLSARSSLAPDSFWAN 241
DB 326 ---VDSNGQVAVSRSRPVAMQGEISVFTKPIAAGVNTITVR----- 368
QY 242 NYSKYAVMGTSNATPIYAAGVAQLRHFIRKNGITPKP---SLIKALILAGATVGL 297
DB 369 --NGYQMGSSQASPTAGAVAVL-----LSAKPASVDALIKALFTSASASA 417
QY 298 GYPSDDQMGKVTLDKSLNAVYVNEATALTITGQKATYSFQIQAGKPLKISLVMTDAPGST 357
DB 418 --KNNNVGFGQISIPGALG---KVGSTGTSQPTPTPPAPPTPTPTPTPTPPAP 469
QY 358 T-----ASYLVNDLVLITAPNGQYV---NDFSPYDNMWDGRNVENVF 402

```

Sat Apr 3 17:20:03 2004

us-09-985-689a-3.rpr

Page 8

Db 470 TPPAPAKPENGPAQYTL-----CAVEGRCGSGARQKQVAFGTDGRWISGTSDDTP 520

Search completed: March 31, 2004, 16:10:17
Job time : 12.3246 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 7.49423 Seconds
(without alignments)
3008.498 Million cell updates/sec

Title: US-09-985-689A-3

Perfect score: 2263
Sequence: 1 NDVARGIKVADYQNNNGLY.....EVQAVNPSPGRRSLAIVH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	502	22.2	1743	1 TAGC_DICDI	Q23668 dictyostei
2	463.5	20.5	1905	1 TAGC_DICDI	P54683 dictyostei
3	332	13.8	806	1 SUBT_BACST	P29141 bacillus su
4	305	13.5	1398	1 PLS_PIRFU	P72186 pyrococcus
5	296.5	13.1	580	1 EXPR_XANCP	P23114 xanthomonas
6	269	11.9	269	1 PRIN_BACSP	Q99405 bacillus sp
7	269	11.9	269	1 SUBS_BACLE	P29600 bacillus le
8	269	11.9	380	1 ELYA_BACAO	P27693 bacillus al
9	269	11.9	380	1 ELYA_BACCS	P43162 bacillus cl
10	265.5	11.7	420	1 SUBT_BACCS	P28842 bacillus sp
11	265	11.7	534	1 PROA_VIBAL	P16588 vibrio algi
12	259	11.4	269	1 SUBB_BACLE	P29599 bacillus le
13	255	11.3	401	1 THEB_BACSP	Q45670 bacillus sp
14	248.5	11.0	1167	1 SCAL_STRPY	P15926 streptococc
15	246	10.9	894	1 WPRB_BACSV	P54423 bacillus su
16	243.5	10.8	378	1 ELYA_BACSP	P20724 bacillus su
17	243	10.7	513	1 AQUJ_THBAQ	P08594 thermus agu
18	242	10.7	379	1 SUBT_BACLI	P00780 bacillus li
19	239.5	10.6	1181	1 SCAL_STRPY	P58059 streptococc
20	239.5	10.5	422	1 TKSU_PYRKO	P29140 bacillus cl
21	237.5	10.5	422	1 ALP_CEPAC	P58502 pyrococcus
22	236.5	10.5	402	1 ALP_CEPAC	P29118 cephalospor
23	231.5	10.2	1052	1 MSIF_HTMAN	P14703 homo sapien
24	229.5	10.1	339	1 ISPI_BACSU	P11018 bacillus su
25	228	10.1	1052	1 MSIF_CRIGR	Q92248 cricetus su
26	228	10.1	1052	1 MSIP_MOUSE	Q9W422 mus musculu
27	227.5	10.1	1052	1 MSIP_RAT	Q9W423 ratius norv
28	227.5	10.1	381	1 SUBN_BACNA	P35835 bacillus su
29	227.5	10.1	381	1 SUBT_BACSA	P00783 bacillus su
30	227.5	10.0	381	1 SUBT_BACST	P29142 bacillus su
31	226.5	10.0	381	1 SUBT_BACST	P29149 bacillus su
32	226.5	10.0	1433	1 SUBT_BACSU	P16397 bacillus su
33	225	9.9	382	1 SUBT_BACAM	P00782 bacillus am

34	224.5	9.9	275	1 SUBT_BACPU	P07518 bacillus pu
35	224	9.9	274	1 SUBD_BACLI	P00781 bacillus li
36	223.5	9.9	279	1 THEB_THBVU	P04072 thermocactin
37	222	9.8	595	1 BPRX_BACNO	P42780 bacteroides
38	220	9.7	404	1 SMP1_MAGPO	Q9Y778 magnaporthe
39	219	9.7	388	1 CUDP_METAN	P29138 metarhizium
40	217.5	9.6	326	1 ISP_PABPO	P29139 paenibacill
41	213.5	9.4	409	1 ALP_TIRHA	Q03420 trichoderma
42	213	9.4	467	1 ISB6_SCHPO	P40903 schizosacch
43	210	9.3	533	1 FBEFC_ASENG	P33295 aspergillus
44	209.5	9.3	451	1 YLPI_SCHPO	Q9UT50 schizosacch
45	209	9.2	530	1 HLY_HAUI7	P29143 halophilic

ALIGNMENTS

RESULT 1
TAGC_DICDI
ID TAGC_DICDI STANDARD; PRT; 1743 AA.
AC Q23668:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Prestalk-specific protein tagc precursor (EC 3.4.21.-).
GN TAGC.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
CX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA MEDLINE=97140317; PubMed=8986798;
RX Shaulsky G., Escalante R., Loomis W.F.;
RT "developmental signal transduction pathways uncovered by genetic
suppressors";
RL Proc. Natl. Acad. Sci. U.S.A. 93:15260-15265(1996).
CC -1- FUNCTION: Inter-cellular communication via tagc may mediate
integration of cellular differentiation with morphogenesis (By
similarity).
CC -1- SIMILARITY: In the N-terminal section; belongs to peptidase family
S8.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ATP-BINDING
TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -1- SIMILARITY: STRONG, TO TAGC.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U60086; AAB0331.1; -
DR PIR; T18279; T18279.
DR DictyBase; DDB0001795; tagc.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001140; ABC_TM_transport.
DR InterPro; IPR003439; ABC_Transporter.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PS00723; SUBTILISIN.
DR SMART; SM00382; AAA_1
DR PROSITE; PS50929; ABC_TMIF; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00136; SUBTILASE_SER; FALSE_NEG.
DR Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane;

```

KW Signal.
FT SIGNAL. 1 27
FT CHAIN 28 1743
FT DOMAIN 316 642
FT DOMAIN 1450 1687
FT TRANSMEM 962 982
FT TRANSMEM 1027 1047
FT TRANSMEM 1072 1092
FT TRANSMEM 1157 1177
FT TRANSMEM 1260 1280
FT TRANSMEM 1286 1308
FT ACT SITE 325 325
FT ACT SITE 372 372
FT ACT SITE 637 637
FT NP BIND 1485 1492
FT DOMAIN 42 46
FT DOMAIN 94 103
FT DOMAIN 643 646
FT DOMAIN 733 741
FT DOMAIN 786 792
FT DOMAIN 1337 1340
FT DOMAIN 1346 1352
FT DOMAIN 1353 1357
FT DOMAIN 1358 1367
FT DOMAIN 1381 1386
FT DOMAIN 1707 1729
FT CARBOHYD 390 390
FT CARBOHYD 536 536
FT CARBOHYD 547 547
FT CARBOHYD 614 614
FT CARBOHYD 689 689
FT CARBOHYD 735 735
FT CARBOHYD 741 741
FT CARBOHYD 766 766
FT CARBOHYD 832 832
FT CARBOHYD 887 887
FT CARBOHYD 1251 1251
FT CARBOHYD 1385 1385
FT CARBOHYD 1386 1386
FT CARBOHYD 1454 1454
FT CARBOHYD 1704 1704
SQ SEQUENCE 1743 AA; 194145 MM; 12DB363E2F729839 CRC64;

Query Match Best Local Similarity 22.2%; Score 502; DB 1; Length 1743;
Matches 162; Conservative 78; Mismatches 158; Indels 190; Gaps 24;

19 LVGGQGVVAVADTGLDGR--NDS-----SMHEAFRGKITALYALGRTNANNDPNC 68
314 LRGGQGLTSLADTGLDGHSCFPDSKYPILANSVNLNHR-KVVTYITTSIDSDSKVDGH 372
69 GTHVAGVVG--NAIN-----KQMAFQANLVQSIMDSGG-GLGPNINLTFSQAWNA 120
373 GTHICGSAAGTPEDSSVNISSPGLADAKIAFDLASSGSLTP-PSDLKQLYQLYLA 431
121 GARLHTNSWGA---PVNGAYTANSRQVDEYV-RNNDMTVLPAAGNEGPNSGTIS--APG 173
432 GARVHCHDSWGSVSVEGYTGSYSDTASIDFLFHPHFILRAAGN---NEQYLSLDTOS 488
174 TAKAIVVGATE-----NYRPSFG----- 192
489 TAKAVITVGAHQTHENYVLTDDPNYINYGSSVDINQELICDPDSRCNVTYTAQCCLESNA 548
193 -----SIAD-----NPNHIAQFSSRGATRDGRKIPDVATGTFILSA 229
549 TTGASCCPTLRLRSVIDAANTQPLNENNICFSSSKGPTHGKMKPALVAGEYITSA 608
230 RSSSLA-----PDSFANYNYSKAMGCTSMAPFYAGVAQLRHH-----FIK 273
609 RSNGANTTDQCGDSSL-ENTVALLA-LSGTSMAISFAAAATTIRQYLVGVIYPTGSIYE 666
274 NRGITPKPSLIKALILGA-----TDVGLGVPSCD-----QSGMRVTLTD 312

```

```

Db 667 SNKQPTGSLKLKALMINNAQLNGTFQLITSSITYPNSQVFNFAASLVOGCAIRMS 726
Qy 313 KSLNVAVNENAT-----ALTTGKATYSF----- 336
Db 727 NMHLVAVNNNSNNNNKTSQGITFEDIGLDLRLYKPNOMKEESLSTGQWTSICFYKRS 786
Qy 337 --OTQAGKPLK--ISLWTDAGSSTTAASTYTLVNDLVLITAPNGQKYGNDPSYDNL- 390
Db 787 SSSNSGNNIPRVAVTLWTDPPSYAGAKENLVNNDLTM-----YRDNSTIYFNSQ 841
Qy 391 -----NWDGRNVNENFINAPQSGTYTEVQAVNPGSGPQRS 428
Db 842 GGSSFFGLAPDTQTLNNVSGIVNPTPEMTYRPMVAGTVNPGPQNS 889

RESULT 2
TAGB_DICDI
ID TAGB_DICDI STANDARD; PRT; 1905 AA.
AC P54683;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Prestalk-specific protein tagb precursor (EC 3.4.21.-).
GN TAGB.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=95262903; PubMed=7744252;
RX STRAIN=AX4;
RA Shaulsky G., Kuapa A., Loomis W.F.;
RT "A multistep resistance transporter/serine protease gene is required for prestalk specialization in Dictyostelium."
RL Genes Dev. 9:1111-1122(1995).
CC -1- FUNCTION: Intercellular communication via tagb may mediate integration of cellular differentiation with morphogenesis.
CC -1- SIMILARITY: In the N-terminal section, belongs to peptidase family S8.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS). MDR SUBFAMILY.
CC -1- SIMILARITY: STRONG, TO TAGC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL, U20432; AAA62212.1; -.
CC PIR: T18267; T18267.
CC MEROPS: S08.10M.
CC DicyBase; DB0001564; tagb.
CC InterPro; IPR003593; AAA_Atpase.
CC InterPro; IPR001140; ABC_TM_transp.
CC InterPro; IPR003439; ABC_transporter.
CC InterPro; IPR002029; Peptidase S8.
CC Pfam; PF00664; ABC_membrane; 1.
CC Pfam; PF00065; ABC_tran; 1.
CC Pfam; PF00082; Peptidase S8; 1.
CC PRINTS; PR00723; SUBTILISTIN.
CC ProDom; PD000006; ABC_transporter; 1.
CC SMART; SM00382; AAA; 1.
CC PROSITE; PS00529; ABC_TM1F; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
CC PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
CC PROSITE; PS00137; SUBTILASE_HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; 1.
CC Hydrolase; Serine protease; ATP-binding; Transport; Transmembrane;
KW Signal.

```

```

FT SIGNAL 1 31 POTENTIAL.
FT CHAIN 32 1905 PROTEASE.
FT DOMAIN 378 700 ABC TRANSPORTER.
FT TRANSMEM 1518 1756 POTENTIAL.
FT TRANSMEM 1011 1031 POTENTIAL.
FT TRANSMEM 1076 1096 POTENTIAL.
FT TRANSMEM 1121 1141 POTENTIAL.
FT TRANSMEM 1210 1230 POTENTIAL.
FT TRANSMEM 1309 1329 POTENTIAL.
FT TRANSMEM 1332 1352 POTENTIAL.
FT ACT_SITE 387 387 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 432 432 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 695 695 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT NP_BIND 1553 1560 ATP (POTENTIAL).
FT DOMAIN 63 67 POLY-GLN.
FT DOMAIN 95 104 POLY-ASN.
FT DOMAIN 107 134 POLY-ASN.
FT DOMAIN 311 321 POLY-SER.
FT DOMAIN 833 837 POLY-SER.
FT DOMAIN 838 844 POLY-GLY.
FT DOMAIN 871 876 POLY-LEU.
FT DOMAIN 1012 1015 POLY-ILE.
FT DOMAIN 1386 1389 POLY-GLU.
FT DOMAIN 1398 1404 POLY-GLY.
FT DOMAIN 1445 1450 POLY-GLY.
FT DOMAIN 1450 1450 POLY-ASN.
FT DOMAIN 1765 1779 POLY-ASN.
FT DOMAIN 1782 1785 POLY-SER.
FT DOMAIN 1807 1812 POLY-PRO.
FT DOMAIN 1813 1860 POLY-GLN.
FT DOMAIN 1872 1878 POLY-PRO.
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 747 747 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 823 823 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1172 1172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1522 1522 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1658 1658 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1905 AA; 212518 MW; B8E223FA8B9AE13C CRC64;

```

```

Query Match 20.5%; Score 463.5; DB 1; Length 1905;
Best Local Similarity 28.0%; Pred. No. 2,6e-22;
Matches 162; Conservative 77; Mismatches 156; Indels 183; Gaps 28;

QY 19 LVGGQGVAVADTGLDTR--NDS-----SMHEAFRGKITLALYALGRTNANDP--N 66
DB 376 LRGKQGLSLADTGLDGHSCFSDSKYPIPNQVNMENHRKVITYI-----TYHDMEDYVN 430
QY 67 GHGTHVAGSVLGN-----ALN--KMAQOANLVQSIIDSSGGGL--PSNLTLFSGA 117
DB 431 GHGTHVCGSAGTPEDDSWAISFSGLATDAXIAFYD--SSGSSEPTTPPEYSGMYRPL 488
QY 118 WNAQAHINNSWGA-----PVNGAYTANSQVDEYV--RNNDMVFLFAAGNEGNSGTSIAP 172
DB 489 YDGAARVHDSWGSVLSQGYGVGSDDAGDIAFLVEYEFELILAAQN--NELFASLLAQ 547
QY 173 GTAANAIVGATE---NT-----RPSF----- 191
DB 548 ATKNAITVGABCTAHVNVSDALEYDESDNANFORPLCFPKKYCNNTTAKCCSEVSNV 607
QY 192 -----GSIAI-----NPNHIAQFSRGATPRGRIRKPEVTAPGTILSAR 230
DB 608 KGIQLCCPASIKMNASDSTTQOPFNENMGFSKGGTHGRKLPDLVAAGEVITBAR 667
QY 231 SS-----LAPDSSEWANTNSKYAYVGTSMATPIYAGVAVOLREHF-----IKN 274
DB 668 SNGENSTDOCGDSL--PNANGILMS--ISGTSMATPLATAATTLIRQLYDVGYPTESEYBE 725
QY 275 RGTPEKPSITIKALLINGADVGLGY-----PSGD-----QGMGRVTLDSL 315
DB 726 NKLLPGLSLIKALMINNAQLINTYFWSASSTNPNSAITEQINGANLIDGMALMANNWL 785
QY 316 NNAVYN-----EAT-----ALTGQKATYSPQTO-----AGK 342

```

```

DB 786 YKSSNPPTPPSPFWIGIGIGIKRQKATEWKEDSLSSGLNKSICYFTYKPSSSSGSGCGGCT 845
QY 343 P-LKISLWTDAPGSTTASVTLVNDL-----VITARN--GQKYVGNDFSYPYDN 390
DB 846 PHLVTLVWTDPPSYSGAKFNLVNLLDLLNSDDSDITITNSGSLQPAKVAP--- 902
QY 391 NMDGRNNVENFVNAPQSGTYTIEVQAVNPSGPPRFS 428
DB 903 --DTLNVEGIIINPTKANNYKFTIAGTNVPIGPQKFS 938

RESULT 3
SUBV_BACSU
ID SUBV_BACSU STANDARD; PRT; 806 AA.
AC P29141;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Minor extracellular protease vpr precursor (EC 3.4.21.-).
GN VPR OR IPA-45R OR BSU38050.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 161-195.
RX MEDLINE=92041574; PubMed=1938892;
RA Stoma A., Rufo G.A. Jr., Theriault K.A., Dwyer M., Wilson S.W.,
RA Pero J.;
RT "Cloning and characterization of the gene for an additional
RT extracellular serine protease of Bacillus subtilis."
RJ J. Bacteriol. 173:6889-6895(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95020537; PubMed=7934828;
RA Glaeser P., Kunze F., Arnaud M., Coudart M.P., Gonzalez W.,
RA Hillo M.F., Ionescu M., Lubochinsky B., Marcelino L., Mosser I.,
RA Pressecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
RA Rapoport G., Danchin A.;
RT "Bacillus subtilis genome project: cloning and sequencing of the 97
RT kb region from 325 degrees to 333 degrees."
RJ Mol. Microbiol. 10:371-384(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Boretto M.G., Bessieres P., Bolotin A., Borchert S.,
RA Bortiss R., Boutsier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Broutier S., Bruchet C.V., Calwell I.F., Cummings N.J., Daniel K.A.,
RA Choi S.K., Codani J.U., Connerton I.F., Cummings N.J., Daniel K.A.,
RA Dauter F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabre C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghin S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haeleth J., Harwood C.R., Henaut A.,
RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Jiric B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kunano M.,
RA Kunita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Potwollik S., Prescott A.M.,
RA Pressecan E., Puig P., Punelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale F.,
RA Sager T., Scallan E., Schleich S., Schroeter R., Scofield F.,
RA Setiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Socokan A., Taccori E., Takagi T., Takahashi H., Takemaru K.,
RA Tatenuchi M., Tanakoshi A., Tanaka T., Terstra P., Tomonori A.,
RA Totsato V., Uchiyama S., Vandenberg M., Vanier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzneger T.,

```

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Darchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*." ;
 RL Nature 390:249-256(1997).
 CC -1- FUNCTION: NOT REQUIRED FOR GROWTH OR SPOULATION.
 CC -1- SUBCELLULAR LOCATION: Secretd.
 CC -1- PTM: PROBABLY UNDERGOES C-TERMINAL PROCESSING OR PROTEOLYSIS.
 CC -1- SIMILARITY: Belongs to peptidase family 88.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; W6590; AAA2881.1; -;
 CC EMBL; X73124; CAA51601.1; -;
 CC EMBL; Z99123; CAA15835.1; -;
 CC PIR; A41341; A41341.
 CC HSSP; P00782; 2SMT.
 CC MEROPS; S08; UPA.
 CC Subtilisin; BG10591; VPR.
 CC InterPro; IPR003137; PA.
 CC InterPro; IPR009020; Peptidase_inhb.
 CC Pfam; PF02225; PA.1.
 CC Pfam; PF00082; Peptidase_S8; 1.
 CC PRINTS; PR00723; SUBTILISIN.
 CC PROSITE; PS00136; SUBTILISIN_1.
 CC PROSITE; PS00137; SUBTILISIN_HIS_1.
 CC PROSITE; PS00138; SUBTILISIN_SER_1.
 CC Hydrolyase; Serine protease; Zymogen; Signal; Complete proteome.
 KW SIGNAL 1 28 POTENTIAL.
 FT PROPEP 29 160
 FT CHAIN 161 806 MINOR EXTRACELLULAR PROTEASE VPR.
 FT ACT_SITE 189 189 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 233 233 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 534 534 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SSQUNCE 806 AA; 85608 MW; P984E3BF08B69DDDCRC64;
 Query Match 13.8%; Score 312; DB 1; Length 806;
 Best Local Similarity 23.4%; Pred. No. 6, 1e-13;
 Matches 129; Conservative 56; Mismatches 156; Indels 210; Gaps 17;
 QY 18 GLYGGQVAVADTGLDGR-----NDSSMHEAFRGKITALVALGRTN 60
 DB 177 GYTGKGIKVALIDTGVVNHHPDLKKNFGQYKGYDFVNDYDKPREPTG----- 224
 QY 61 NNDPNG-----HGHYAGSVLGNALNKGMAPQANLVFQSIMSSGGLGSPENLTLFSQ 116
 DB 225 ---DPGEATDGHGTHAGVAAAGITKGVAPDPTLLAYVLDPPG--STTENVAIGVER 279
 QY 117 AMNAGARIHTNSGAEVNGAYTANSPQVDEYVRNNDMTVLPAQNEGSPNGTISAPGTAK 176
 DB 280 AVQDGDADVMNLSLGNSLNPNPDATSTALD-WMASEGVAVVTNGNSGPGWGTVSGPSR 338
 QY 177 NATVGATE---NYPSPFGST----- 194
 DB 339 EALISVATQPLNEVAVTFSYSASAKVWGYNKEDVYKALNKEVELVEAGDEANDFEKG 398
 QY 195 -----ADNP----- 198
 DB 399 DLTKGVAIVKRGSLAFVDKADNAKGAIGMYYNNLSGEIEANVPGMSVPTIKSLDGG 458
 QY 199 -----NHIAQFSSKATRP-GIKEDVTAAPGFTILSAGS 232
 DB 459 EKUVALKAGETKTKTLTVSKALGEOVADFSRRGVMDTWIKDKISAPGVNIVSTIPT 518
 QY 233 LAPDSSFWMYNSKYAVMGTSNATPIVAGNVAQLREHFIKRGITTPKSL--IKALIA 290

DB 519 HDPDHPY--GYGSKQ-----GTSWASPHIAGAVAVIKQ-----AKPKMSVEQIKALIMN 565
 QY 291 GATDV-----GLGYPSGDQGWGRVTLDKSLNAVAYNEATLTGQKATY-SFOTQACKPKK 345
 DB 566 TAVTLKDSDEGVYPHNAQAG-----SARINMAIRKADSLVSPGYSYGFLENGENETK 619
 QY 346 ISLWTDAPSTASATLVNDLVLITAPRGQYVNDSPYDNNWMDGNVENVFIVA 405
 DB 620 NEFTTINQSIKSKSYTL-----EYSPNGSIGISTGTSRVIVPA 658
 QY 406 PQSGTYTIEVQ 416
 DB 659 HQTKATATKVK 669
 RESULT 4
 PLS_PYRPU STANDARD; PRT; 1398 AA.
 ID PLS_PYRPU
 AC P72186;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Pyrolysin precursor (EC 3.4.21.-).
 GN PLS OR PF0287.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 CC Pyrococcus.
 OX NCBI_Taxid=2261;
 RN [1]
 RP SEQUENCE FROM N.A. SEQUENCE OF 150-184, AND CHARACTERIZATION.
 RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RX MEDLINE=96355370; PubMed=8702780;
 RA Voorthorst W.G.B., Eggen R.I.L., Geertling A.C.M., Platteuw C.,
 RA Slezten R.J., de Vos W.M.;
 RT "Isolation and characterization of the hyperthermostable serine
 RT protease, pyrolysin, and its gene from the hyperthermophilic archaeon
 RT *Pyrococcus furiosus*." ;
 RL J. Biol. Chem. 271:20426-20431(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RX MEDLINE=21073021; PubMed=11210516;
 RA de Vos W.M., Voorthorst W.G.B., Dijkgraaf M., Kluskens L.D.,
 RA Van der Oost J., Slezten R.J.;
 RT "Purification, characterization, and molecular modeling of pyrolysin
 RT and other extracellular thermostable serine proteases from
 RT hyperthermophilic microorganisms." ;
 RL Mehl. Enzymol. 330:383-393(2001).
 RN [3]
 RP FUNCTION: Has endopeptidase activity toward caseins, casein
 CC fragments including alpha-SI-casein and synthetic peptides.
 CC -1- SUBCELLULAR LOCATION: Cell envelope associated.
 CC -1- PTM: LMM pyrolysin seems to be produced by autoproteolytic
 CC activation of HMW pyrolysin.
 CC -1- PM: Glycosylated.
 CC -1- MISCELLANEOUS: Thermostable; high activity at 95 degrees Celsius.
 CC -1- SIMILARITY: Belongs to peptidase family 88.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U55835; AAB09761.1; -;
 DR EMBL; AE01053; AAL80411.1; -;
 DR PIR; T28159; T28159.

DR HSSP: Q45670; 1DBI.
 DR MEROPS; S08.100; -;
 DR InterPro; IPR000209; Peptidase_S8.
 DR InterPro; IPR007280; PPC.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR Pfam; PF04151; PPC; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
 KW Complete proteome.
 FT SIGNAL 1 26
 FT PROPEP 27 149
 FT CHAIN 150 1398
 FT ACT_SITE 179 179
 FT ACT_SITE 365 365
 FT ACT_SITE 590 590
 FT CARBOHYD 152 152
 FT CARBOHYD 222 222
 FT CARBOHYD 228 228
 FT CARBOHYD 240 240
 FT CARBOHYD 257 257
 FT CARBOHYD 262 262
 FT CARBOHYD 298 298
 FT CARBOHYD 327 327
 FT CARBOHYD 406 406
 FT CARBOHYD 651 651
 FT CARBOHYD 663 663
 FT CARBOHYD 739 739
 FT CARBOHYD 792 792
 FT CARBOHYD 893 893
 FT CARBOHYD 908 908
 FT CARBOHYD 917 917
 FT CARBOHYD 929 929
 FT CARBOHYD 1048 1048
 FT CARBOHYD 1056 1056
 FT CARBOHYD 1084 1084
 FT CARBOHYD 1117 1117
 FT CARBOHYD 1140 1140
 FT CARBOHYD 1148 1148
 FT CARBOHYD 1208 1208
 FT CARBOHYD 1233 1233
 FT CARBOHYD 1237 1237
 FT CARBOHYD 1332 1332
 FT CONFLICT 607 609
 FT CONFLICT 881 881
 SQ SEQUENCE 1398 AA; 154474 MW; 355D873A27D56552 CRC64;
 Query Match 13.5%; Score 305; DB 1; Length 1398;
 Best Local Similarity 28.1%; Fred. No. 3.5e-12;
 Matches 128; Conservative 37; Mismatches 154; Indels 136; Gaps 15;
 21 GCGQVAVADTGTGDTGRNDS-----SMHEAFRGKITATLALGRTNANDPN-- 66
 301 GNGVDIAVDTDDYPTDEVPVLCQVNTVYDAVVFYVYGPAN--YVLAET----DPNCE 354
 67 -----GHGTVAGSVTGNALN-----KG 84
 355 YAVFGMDGHGHGTVAGTVAGYDSNNDAMDWLSMYSGEWEVSRILYGMVYNTTDTVVG 414
 85 MAPQANLVFQSIDSSGGJGGLPSNLTFLSCAMNAGARLHTNWSG--APVNGAVTANR 142
 415 VARGAGIMAIRVLRSDG--RGSMDIIEGWTVAATGADVISMISLGSNAPYLDGDPREV 472
 143 QVDEYVRANDMTVLFAAGNENGSGTISAPGTAKNAITVGAET----- 185
 473 AVDELTEKVGAVVIAAGNGPGITVSGPVATKXITVGAAPVTVNGVVSQALGYPD 532
 186 ----NVAESPSTADNHNHIAQFSSRGATDGRKIXDVAAPGFTILASASSIAPDSSFWA 241
 533 YVGFYVFPATYV-----RIAFSSRGPRIDGELKENVVAPGIGITSSLPWIMGADF-- 585

QY 242 NYSKAVNGCTSMATPIYAGNTAQLREHFIRKRGITTPKPSLIKALLINGAT-----DV 295
 DB 586 -----MSGTSMATPHVSGVALLISG-AKREGVYNPDITKYLEGATWLEGDPT 636
 QY 296 GLVPSGDQGMGRVTLDDKSLNVAVYNEATLTGQATVSPFOQAKPKLISLVWTDAG 355
 DB 637 GQKTELDQGHGLVNTKSWET-----LKAINGTTLPIYDHMDKRY 678
 QY 356 STTASVTLVNDLDTVTPANG-----QKRYGN 392
 DB 679 SDFAEYLGV-DVIRGLYARNSIPDIYEMHIKYGD 712
 RESULT 5
 EXPR_XANCP STANDARD; PRT; 580 AA.
 ID AC P2314;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Extracellular protease precursor (EC 3.4.21.-).
 GN XCC0851.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_Taxid=340;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90251253; PubMed=2187155;
 RA Liu Y.-N., Tang J.-L., Clarke B.R., Dow J.M., Daniels M.J.;
 RT "A multipurpose broad host range cloning vector and its use to
 characterise an extracellular protease gene of Xanthomonas campestris
 pathovar campestris";
 RL Mol. Gen. Genet. 220:433-440(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=22022145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., de Azeiteiro A.M., Bertolini M.C., Camargo L.E.A.,
 RA Camarotte G., Camaran F., Cardoso J., Chamberg F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
 RA Fortuglieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locali E.C., Machado M.A., Madeco A.M.B.N., Martinez-Rossi N.M.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.P.,
 RA Spinoia L.A.F., Takita W.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Seubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 host specificities";
 RL Nature 417:459-463(2002).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to peptidase family 58.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC DR EMBL; X51635; CA35962.1; -;
 DR EMBL; AE012184; AA040166.1; -;
 DR PIR; S11890; S11890.
 DR HSSP; P00782; 28BT.
 DR MEROPS; S08.TPA; -;

```

DR InterPro: IPR000209; Peptidase_S8.
DR InterPro: IPR007280; PPC.
DR InterPro: IPR009020; Protease_inhib.
DR Pfam: PF00082; Peptidase_S8; 1.
DR Pfam: PF04151; PPC; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR HydroLase; Serine protease; Zymogen; Signal; Complete proteome.
FT SIGNAL 1 32
FT PROPEP 33 32
FT CHAIN 7137 580
FT ACT_SITE 177 177
FT ACT_SITE 237 237
FT ACT_SITE 409 409
FT DISULFID 225 273
FT DISULFID 315 352
FT DISULFID 450 454
SQ SEQUENCE 580 AA; 57228 MW; 8C9A2CBE4E7F47CB CRC64;

Query Match 13.1%; Score 296.5; DB 1; Length 580;
Best Local Similarity 29.0%; Pred. No. 4.1e-12;
Matches 137; Conservative 52; Mismatches 148; Indels 135; Gaps 26;

QY 21 GGGVVAVVADTGL---DTGRNDSMHEAFRGKITLVLGRTNNAND----- 64
DB 168 GSGTVAVVADTGLTTSADLNANILAGYDPLSDATTRDNGSDSNADGDMVYANNEGA 227
QY 65 -----FNGGTHVAGS-----VLGNALNKGMAFOANL-----VFQSIMDS 99
DB 228 GIPAASSSMHGTHVAGTVAAVTNNTTGAAGVAVRVYLGKCGSLDIADAIWA 287
QY 100 SGG-LEGLPSNUN--TLFQAMNAGARHTNSMGAPVNAAYANRQVDYIRANDMYL 156
DB 288 SGGTVSGIPANAPAVINNSLGGGSCSTWONA--INAVSRGT-----TVV 334
QY 157 FAAGNCGPN-SGTISAPGTAKNAITGATENYRPSFGSIADPNHIAQFSSRGATRDGRI 215
DB 335 VAAGNDAENVSG--SLPANCAVIAVAAT-----TSAGA-----KASYSNGT----- 375
QY 216 KPDVTAAGTITLSARSS--LAPDSSFWANYSKYAMGTSMTATPVAGNVAQLREHFIR 273
DB 376 GIDVSAFGSILSTLNSGTTTSGSASVAYN-----GTSMASPHVAAGVALVOS--VA 426
QY 274 NRGITPK--PSLIK--AALIAGATDVGLG-----PSGDQMGRTVLDSK 314
DB 427 PTLATPAAVETLLKNTARALPGACSGCGAGIVNAAATAALINGSGGGGGGNTLTNG 486
QY 315 LNVAYVNEATALTGQKATYSFQTQNGKPLKISLVTPDAPSGSTTASVTLVNDLVI--- 371
DB 487 TPVTGGA---TGAEVNTITIVPAGSG--TLVTTTSGS-----GADIYVAG 531
QY 372 TAENGKXYVNDPSYPDNMNGNNVENVFINAPSGTYTTEVOAYNVPSG 423
DB 532 SAFTDSAYT---CRPYRS-----GNAETCTIAP-SGTYVRLKXAYSTPSG 573

```

```

RT "Structure of a new alkaline serine protease (M-protease) from
RT Bacillus sp. KSM-X16."
RL Acta Crystallogr. D 51:199-206(1995).
RN [2]
RP SEQUENCE OF 1-23, AND CHARACTERIZATION.
RX MEDLINE=95358832; PubMed=763397;
RA Kobayashi T., Hakamada Y., Adachi S., Hitomi J., Yoshinatsu T.,
RT "Purification and properties of an alkaline protease from
RT alkalophilic Bacillus sp. KSM-X16."
RL Appl. Microbiol. Biotechnol. 43:473-481(1995).
CC -1- COPACTOR. Binds 2 calcium ions per subunit.
CC -1- SUBUNIT: Monomer.
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC PDB; 1MPT; 22-JUN-94.
DR MEROPS; S08.010; -.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DX HydroLase; Serine protease; Metal-binding; Calcium-binding;
DX 3D-structure.
FT ACT_SITE 32 32
FT ACT_SITE 62 62
FT ACT_SITE 215 215
FT METAL 2 2
FT METAL 40 40
FT METAL 73 73
FT METAL 75 75
FT METAL 77 77
FT METAL 79 79
FT METAL 163 163
FT METAL 165 165
FT METAL 168 168
FT HELIX 6 10
FT TURN 11 12
FT TURN 13 18
FT TURN 19 20
FT TURN 24 25
FT TURN 27 32
FT TURN 39 40
FT TURN 43 48
FT TURN 51 52
FT HELIX 62 71
FT TURN 84 85
FT TURN 87 92
FT TURN 96 97
FT HELIX 102 114
FT TURN 115 116
FT STRAND 119 122
FT STRAND 126 126
FT HELIX 131 142
FT TURN 143 144
FT STRAND 146 150
FT STRAND 161 161
FT TURN 162 164
FT TURN 166 167
FT STRAND 169 174
FT TURN 176 177
FT STRAND 180 180
FT TURN 182 183
FT TURN 188 189
FT STRAND 192 195
FT STRAND 199 203
FT TURN 204 206
FT STRAND 207 211
FT HELIX 214 231
FT TURN 233 234
FT HELIX 237 246
FT TURN 247 249

```

FT HELIX 254 257
 FT TURN 258 259
 FT STRAND 261 261
 FT HELIX 264 267
 FT TURN 268 269
 SQ SEQUENCE 269 AA; 26723 MW; 7A03C86D534A1D07 CRC64;

Query Match 11.9%; Score 269; DB 1; Length 269;
 Best Local Similarity 31.7%; Pred. No. 9,3e-11;
 Matches 98; Conservative 32; Mismatches 97; Indels 82; Gaps 14;

QY 13 AANNVGLYQGGQVAVADTGLDGRNDSMHEAFKRTALYALGRNNANDPNGHGTHV 72
 DB 15 AAHNRGLTGSQVAVLDTGIST-----HPDLNIRGASFPVGGP-STQGNHGTHV 66
 QY 73 AGSV--LGNALN-KGMAPQANLVFQSIMSSGGLGSPNNTLFSQANNAARHTHTSM 129
 DB 67 AGTIALNNSIGVLGAPSAELYAVKVLASG--SGSVSSIQGLEWAGNNGMHVNLSL 124
 QY 130 GAPVNGAYTANSRQVDEYVRNNDMTVLFPAENGEPNSGTISAPGTAKNAITVGATENYRP 189
 DB 125 GSP---SPSATELGAVNSATSRGLVVAASGNSG--AGSISYPARYANMAVGAT----- 174
 QY 190 SFGSINDNPHTAQSSRGATRDGRIRKPVTAAGTIFILARSGLAPDSFTANYSKAY 249
 DB 175 -----DQNNRASFQYGAGL-----DIVAPGVAVQSTYP-----GSTVAS 210
 QY 250 MGGTSNATPIVAGNVA-----OLREHFIRNGITPKPSLIKALTAGATDVL 297
 DB 211 LNTGSMATPHVAGVALVVKQKPSMSNGVRNH-LKN-----TANGL 251
 QY 298 G---YPSG 302
 DB 252 GNTNLVSGS 260

RESULT 7
 SUBS_BACLE STANDARD; PRT; 269 AA.
 AC P29600;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Subtilisin Savinase (EC 3.4.21.62) (Alkaline protease).
 OS Bacillus lentus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1467;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
 RX MEDLINE=92148829; PubMed=1738156;
 RA Betzel C., Klupsch S., Papendorf G., Hastrop S., Branner S.,
 RA Wilson K.S.;
 RT "Crystal structure of the alkaline proteinase Savinase from Bacillus
 RT lentus at 1.4-A resolution.";
 RT J. Mol. Biol. 223:427-445(1992).
 RN [2]
 RP STRUCTURE BY NMR.
 RX MEDLINE=96184541; PubMed=8654411;
 RA Remerowski M.L., Pepermans H.A.M., Hilbers C.W., van de Ven F.J.M.;
 RT "Backbone dynamics of the 269-residue, protease Savinase determined
 RT from 15N-NMR relaxation measurements.";
 RT Eur. J. Biochem. 235:629-640(1996).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (0.78 ANGSTROMS).
 RX MEDLINE=9846039; PubMed=9753430;
 RA Kuhn P., Knapp M., Soltis S.M., Ganshaw G., Thoenes M., Bort R.;
 RT "The 0.78-A structure of a serine protease: Bacillus lentus
 RT subtilisin.";
 RT Biochemistry 37:13446-13452(1998).
 CC -!- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
 CC it catalyzes the hydrolysis of proteins and peptide amides.
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
 CC for peptide bonds, and a preference for a large uncharged residue

CC in PL. Hydrolyzes peptide amides.
 CC -!- COFACTOR: Binds 2 calcium ions per subunit.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- BIOTECHNOLOGY: Used as a detergent protease. Sold under the name
 CC Savinase by Novozymes.
 CC -!- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
 CC sporulation, and many mutations which block sporulation at early
 CC stages affect expression levels of subtilisin. However, subtilisin
 CC is not necessary for normal sporulation.
 CC -!- SIMILARITY: Belongs to peptidase family S8.
 DR PDB; 1C9J; 06-OCT-99.
 DR PDB; 1C9M; 10-JAN-01.
 DR PDB; 1C9N; 10-JAN-01.
 DR PDB; 1GCI; 11-NOV-98.
 DR PDB; 1IAY; 11-JUL-01.
 DR PDB; 1FEA; 26-NOV-97.
 DR PDB; 1SVN; 14-OCT-96.
 DR MEROPS; S08.103; -.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PRO0723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR HydroLase; Sporulation; Serine protease; Metal-binding;
 KW Calcium-binding; 3D-structure.
 FT ACET SITE 32 32
 FT ACET SITE 62 62
 FT ACET SITE 215 215
 FT METAL 2 2
 FT METAL 40 40
 FT METAL 73 73
 FT METAL 75 75
 FT METAL 77 77
 FT METAL 79 79
 FT METAL 163 163
 FT METAL 165 165
 FT METAL 168 168
 FT STRAND 2 2
 FT HELIX 6 6
 FT TURN 11 11
 FT HELIX 13 13
 FT TURN 19 20
 FT TURN 24 25
 FT STRAND 27 32
 FT TURN 39 40
 FT STRAND 43 48
 FT TURN 51 52
 FT HELIX 62 71
 FT STRAND 78 78
 FT TURN 84 85
 FT STRAND 87 92
 FT STRAND 94 94
 FT TURN 96 97
 FT STRAND 100 100
 FT HELIX 102 114
 FT TURN 115 116
 FT STRAND 119 122
 FT STRAND 126 126
 FT HELIX 131 142
 FT TURN 143 144
 FT STRAND 146 150
 FT STRAND 161 161
 FT TURN 162 164
 FT TURN 166 167
 FT STRAND 168 174
 FT TURN 176 177
 FT STRAND 180 180
 FT TURN 182 183
 FT TURN 188 189
 FT STRAND 190 195
 FT TURN 199 203
 FT TURN 204 206

```

FT STRAND 207 211
FT HELIX 214 231
FT TURN 233 234
FT HELIX 237 246
FT TURN 247 247
FT STRAND 249 249
FT HELIX 254 257
FT TURN 258 259
FT STRAND 261 261
FT HELIX 264 267
FT TURN 268 269
SQ SEQUENCE 269 AA; 26698 MW; 4D89F8778999BE8D CRC64;

Query Match 11.8%; Score 269; DB 1; Length 269;
Best Local Similarity 31.8%; Pred. No. 9.3e-11;
Matches 97; Conservative 35; Mismatches 99; Indels 74; Gaps 14;

QY 13 AONNYGLYGGGVAVADTGLDGRNDSMHEAFKRTALVALGRTNNANDPNHGTHV 72
Db 15 AANNRGITSGVAVADTGLDGLTST-----HPDLNIRGASFPVPSGP-STOGNGHGTHV 66
QY 73 AGSV--LGNALN-KGNAPQANLVFOSIMDSGGIGGLPSNLTIFSOAMNAGARIHTNSM 129
Db 67 AGTIALNNISIGVIGVAPSAHELYAVKLGASG--SGSVASIQGLEMAGNNGHVALNLSL 124
QY 130 GAVVNGAYTANSGVDEYVRNNMTVLPAAANGNGRSGTISAPGTAKAITYGATENTRP 189
Db 125 GSP--SPSATLEQAVNSATSRGVLVVAASGNSG--AGSISYFARYANMAVAGAT----- 174
QY 190 SFGSIADNPNHIAQFSSRGATRDGRIPDVTAPGTFLIARSLSLADSSFWANYSKYAY 249
Db 175 -----DONNRRASFQYGAGL-----DIVAPGVAVQSTYP-----GSTYAS 210
QY 250 MGGTSWATPVAAGNVA-----QIREFTNRGRTPKPSLIKALTAGATDVL 297
Db 211 LNSTSATPHTVAGAAALVKQKNSMSNVQIRNH-LNN-----YATSLGSTNL-- 256
QY 298 GYPSG 302
Db 257 -YGSQ 260

RESULT 8
ELYA_BACAO
ID ELYA_BACAO STANDARD; PRT; 380 AA.
AC P27693;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alkaline protease precursor (EC 3.4.21.-).
OS Bacillus alcalophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1445;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=PB92;
RC MEDLINE=91262483; PubMed=2059048;
RA van der Laan U.C., Gerritse G., Mulleners L.J.M., van der Hoek R.A.,
RA Quax W.J.;
RT "Cloning, characterization, and multiple chromosomal integration of a
RT Bacillus alkaline protease gene.";
RL Appl. Environ. Microbiol. 57:901-909 (1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
RC STRAIN=PB92;
RX MEDLINE=92390330; PubMed=1518788;
RA van der Laan J.C., Teplyakov A.V., Keiders H., Kalk K.H., Misset O.,
RA Mulleners L.J.M., Dijkstra B.W.;
RT "Crystal structure of the high-alkaline serine protease PB92 from
RT Bacillus alcalophilus.";
RL Protein Eng. 5:405-411 (1992).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).

```

```

RX MEDLINE=93078250; PubMed=1447775;
RA Sobek H., Hecht H.-U., Aehle M., Schomburg D.;
RT "X-ray structure determination and comparison of two crystal forms of
RT a variant (Asn154Arg) of the alkaline protease from Bacillus
RT alcalophilus refined at 1.85-A resolution.";
RL J. Mol. Biol. 228:108-117 (1992).
RN [4]
RP STRUCTURE BY NMR OF 112-380.
RX MEDLINE=97277237; PubMed=9115441;
RA Martin U.R., Mulder P.A., Karim-Nejad Y., van der Zwan J.,
RA Martini M., Schipper D., Boelens R.;
RT "The solution structure of serine protease PB92 from Bacillus
RT alcalophilus presents a rigid fold with a flexible substrate-binding
RT site.";
RL Structure 5:521-532 (1997).
CC - CORACOR: Binds 2 calcium ions per subunit.
CC - SUBCELLULAR LOCATION: Secreted.
CC - SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M65086; AAA2212.1; -;
DR EMBL; A13739; CA01126.1; -;
DR PIR; A49778; A49778.
DR PDB; 1AH2; 1S-APR-98.
DR MEROPS; S08.038; -;
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR003020; Protease_inhib.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00135; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR KW Hydrolyase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
KW Signal; 3D-structure.
FT SIGNAL 1 27
FT PROPEP 28 112
FT CHAIN 113 380
FT ACT_SITE 143 143
FT ACT_SITE 173 173
FT ACT_SITE 326 326
FT METAL 113 113
FT METAL 151 151
FT METAL 184 184
FT METAL 186 186
FT METAL 188 188
FT METAL 190 190
FT METAL 274 274
FT METAL 276 276
FT METAL 279 279
FT METAL 297 297
FT TURN 121 123
FT TURN 125 130
FT HELIX 125 130
FT TURN 131 131
FT TURN 133 136
FT STRAND 138 142
FT TURN 150 151
FT STRAND 156 157
FT HELIX 173 182
FT STRAND 198 201
FT TURN 207 208
FT HELIX 213 226
FT STRAND 230 233
FT HELIX 242 252
FT TURN 253 255
FT STRAND 257 258

```


QY 298 GYPSG 302
DB 368 -YSGG 371

RESULT 10
SUBT_BACS9
ID SUBT_BACS9 STANDARD; PRT; 420 AA.
AC P28842;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Subtilisin precursor (EC 3.4.21.62).
GN SUBT.
OS Bacillus sp. (strain TA39).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=29336;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92256481; PubMed=1581352;
RA Narinx E., Davail S., Feller G., Gerday C.;
RT "Nucleotide and derived amino acid sequence of the subtilisin from
the anaerobic psychrotroph Bacillus TA39.";
RL Biochim. Biophys. Acta 1131:111-113 (1992).
CC -1- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
it catalyzes the hydrolysis of proteins with broad specificity
CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins with broad specificity
CC for peptide bonds, and a preference for a large unchanged residue
CC in P1. Hydrolyzes peptide amides.
CC -1- COFACTOR: Binds 1 calcium ion per subunit (potential).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: Still active at temperatures close to 0 degrees
CC Celsius, it has a marked heat lability.
CC -1- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
CC sporulation, and many mutations which block sporulation at early
CC stages affect expression levels of subtilisin. However, subtilisin
CC is not necessary for normal sporulation.
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X62369; CAA44227.1; -
CC PIR: S23407; S23407.
CC HSSP: Q99405; IMPT.
CC
CC MEROPS: S08.004; IMPT.
CC InterPro: IPR000209; Peptidase_S8.
CC Pfam: PF00082; Peptidase_S8; 1.
CC PRINTS: PR00723; SUBTILISIN.
CC PROSITE: PS00136; SUBTILISIN_ASP; 1.
CC PROSITE: PS00137; SUBTILISIN_HIS; 1.
CC PROSITE: PS00138; SUBTILISIN_SER; 1.
CC Hydrolyase; Sporulation; Serine protease; Zymogen; Metal-binding;
KW Hydrolyase; Sporulation; Serine protease; Zymogen; Metal-binding;
KW Calcium-binding; Signal.
FT SIGNAL 1 26
FT PROPEP 27 111
FT CHAIN 112 420
FT ACT_SITE 145 145
FT ACT_SITE 182 182
FT ACT_SITE 360 360
FT METAL 115 115
FT METAL 154 154
SQ SEQUENCE 420 AA; 44086 MW; AE4F12BD3B26EC CRC64;

Query Match 11.7%; Score 265.5; DB 1; Length 420;
Best Local Similarity 29.0%; Pred. No. 2,6e-10;
Matches 92; Conservative 46; Mismatches 120; Indels 59; Gaps 15;

QY 21 GGGVAVVADTGLDTGRNDSMEAFRGKITAL--YALGRT--NNANDPNHGTHVAGS 75
DB 136 GGGINIAVLDTGVTN-----HPDLRNNVGEQCKDFVTGTTNNNCTDRQHGCTHVAS 189
QY 76 VL-----GNALNKGMAQALVFOQSI--DSSGLGLGSPENLNTLFSQANAGARLHTN- 127
DB 190 ALADGTGNGV-YGVAADADLMAYKVLGDGSGYADDIAAIIHADQATALNTKXVIM 248
QY 128 SWGAPVNGAYTANRQYDEYVRANDMTVLPAANGSPNSGTTISAPGTAKNAITVGATENY 187
DB 249 SLSSGSSSLITNA--VNSYKGVLIILAAAGNSGPYQSGICYPGALVNAVVALEN- 304
QY 188 RPSFGSIADNPNNHIAQSSRFG-----ATRDGRKPDVTAPGTTILSARSLADDSSE 239
DB 305 -----KYENGTYRVADPSRSGYSMTDGDYIQKGDV--ELISAPGAAYST----- 347
QY 240 WANYNSKYAVMGSTMATP-VAGNVAQLREHFIKNRGIFKPELTKAALTAGATDVLGY 299
DB 348 W--FDGGYATISGTSMAAPRAAGLAKIMAQYPSANVDYRGLQYRAY---ENDILSGY 402
QY 300 PSG-----DQWGRVTL 311
DB 403 YAGYGDDFASGFGFATV 419

RESULT 11
PROA_VIBAL
ID PROA_VIBAL STANDARD; PRT; 534 AA.
AC P16588;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alkaline serine exoprotease A precursor (EC 3.4.21.-).
GN PROA.
OS Vibrio alginolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=663;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8326126; PubMed=2546861;
RA Deane S.M., Robb F.T., Robb S.M., Woods D.R.;
RT "Nucleotide sequence of the Vibrio alginolyticus calcium-dependent,
RT detergent-resistant alkaline serine exoprotease A.";
RL Gene 76:281-288 (1989).
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M25499; AAA27550.1; -
CC PIR: JS0173; JS0173.
CC HSSP: Q99405; IMPT.
CC
CC MEROPS: S08.050; IMPT.
CC InterPro: IPR000209; Peptidase_S8.
CC InterPro: IPR007280; PPC.
CC Pfam: PF00082; Peptidase_S8; 1.
CC PRINTS: PR00723; SUBTILISIN.
CC PROSITE: PS00136; SUBTILISIN_ASP; 1.
CC PROSITE: PS00137; SUBTILISIN_HIS; 1.
CC PROSITE: PS00138; SUBTILISIN_SER; 1.
CC Hydrolyase; Serine protease; Zymogen; Signal.
KW Hydrolyase; Serine protease; Zymogen; Signal.
FT SIGNAL 1 21
FT PROPEP 22 141
FT CHAIN 142 534
SQ SEQUENCE 534 AA; 534 MW; ALKALINE SERINE EXOPROTEASE A.

FT ACT_SITE 180 180 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 213 213 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT_SITE 363 363 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 534 AA, 55930 MW, 84896DC649DA226 CRC64;

Query Match 11.7%; Score 265; DB 1; Length 534;
 Best Local Similarity 26.3%; Pred. No. 4, 1e-10; Indels 128; Gaps 24;
 Matches 123; Conservative 65; Mismatches 152;

QY 7 IVKADVAQ-----NNY--GLYGGQGVAVADTGLDTRGDSMHRAP 46
 DB 137 IVSADANQNMALMGDRIDRQRLPLDNNNSANPDDGVTAVYIDGV-----NNAHAF 190
 QY 47 RGRITALLY-ALGRTNNANDPNNGHTVAGSVLGNALNKMAQANLVQSITM--DSGGL 103
 DB 191 GGRSVSGVDFVNDADASDCNGHGHVAGTIGGSILY--GVAANVNLVGVTVLSCSGSGST 248
 QY 104 GGLPSNLNTLFSQANAGARHTNSWGAFVNGAVYANSRQVDEYVRNNDMTVLFAAGNEG 163
 DB 249 SGVIAGVDVVAANA--SGFSVANMSIG---GGQSVALDAVQSAVQSG--VSTMLAAGNSN 302
 QY 164 PMSGTISAPGTAKNAITVGATENYRPSFGSIADNPENHIAQSSRGATRDGRIKPDVYAPG 223
 DB 303 ADACNYS-PARVATGVTVGST-----TSTDARSSFSNMQSC-----VDVFAPG 344
 QY 224 TFLISARSSLAPDSSFMANYNSKYAMGGSMTATPIVAGNVQLEHFIKNGRITPKPSL 283
 DB 345 SQTLSA-----W--YDGGYKTISGTSMTATPVAG-VALL--YLOENSVS--PSQ 387
 QY 284 IKAALIGA-----TDVGLGYPSG-----DQMGRTVLDSKLNAY 319
 DB 388 VERLIVSRASTGKVTDRGSVNKLIVSLTDADCGQCGGPDPTPPECKLTGCVVS--- 444
 QY 320 VNEATLTLTGQKATYFQQAQKPLKLSLWTDAPSTASTASTIVNDLVIYRANGQX 379
 DB 445 ---GLSSSGQVAYVYDVVAGQRLTVQW-----YGGSGDADLYRFP--GAK- 486
 QY 380 VGNDPSYPYDNWMDGR---NNVENVFINAPOSGITLIEVQAYNPVSG 423
 DB 487 -----PILNAMDGRPFKYGNNETCTGATGSGRHHVMIQGYSNVSG 527

RESULT 12
 SUBB_BACLE STANDARD; PRT; 269 AA.
 ID SUBB_BACLE STANDARD; PRT; 269 AA.
 AC P29599;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Subtilisin BL (EC 3.4.21.62) (Alkaline protease).
 OS Bacillus lentus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1467;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (1.4 ANGSTROMS).
 RA MEDLINE=93085738; PubMed=1453465;
 RA Gedeckte D.W., Paech C., Yang S.S., Melenz J.R., Bystroff C.,
 RA "The crystal structure of the Bacillus lentus alkaline protease,
 RT subtilisin BL, at 1.4-A resolution.";
 RT J. Mol. Biol. 228:580-595 (1992).
 CC -1- FUNCTION: Subtilisin is an extracellular alkaline serine protease,
 CC it catalyzes the hydrolysis of proteins and peptide amides
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of proteins and peptide amides
 CC for peptide bonds, and a preference for a large unchanged residue
 CC in P1. Hydrolyzes peptide amides.
 CC -1- COFACTOR: Binds 2 calcium ions per subunit.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: Secretion of subtilisin is associated with onset of
 CC sporulation, and many mutations which block sporulation at early
 CC stages affect expression levels of subtilisin. However, subtilisin
 CC is not necessary for normal sporulation.
 CC -1- SIMILARITY: Belongs to peptidase family 58.

DR PDB; 1ST3; 31-JAN-94.
 DR MEROPS; S08.003; -.
 DR InterPro; IPR00209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00136; SUBTILASE_SER; 1.
 DR PROSITE; PS00136; SUBTILASE_1.
 DR Hydrolyase; Sporulation; Serine protease; Metal-binding;
 KM Calcium-binding; 3D-structure.
 QY ACT_SITE 32 32 CHARGE RELAY SYSTEM.
 FT ACT_SITE 62 62 CHARGE RELAY SYSTEM.
 FT ACT_SITE 215 215 CHARGE RELAY SYSTEM.
 FT METAL 2 2 CALCIUM 1.
 FT METAL 40 40 CALCIUM 1.
 FT METAL 73 73 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 75 75 CALCIUM 1.
 FT METAL 77 77 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 79 79 CALCIUM 1 (VIA CARBONYL OXYGEN).
 FT METAL 163 163 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 165 165 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 168 168 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 6 6 CALCIUM 2 (VIA CARBONYL OXYGEN).
 FT METAL 11 12 CALCIUM 1.
 FT METAL 13 18 CALCIUM 1.
 FT METAL 19 20 CALCIUM 1.
 FT METAL 24 25 CALCIUM 1.
 FT METAL 27 32 CALCIUM 1.
 FT METAL 39 40 CALCIUM 1.
 FT METAL 43 48 CALCIUM 1.
 FT METAL 51 52 CALCIUM 1.
 FT METAL 62 71 CALCIUM 1.
 FT METAL 84 85 CALCIUM 1.
 FT METAL 87 92 CALCIUM 1.
 FT METAL 96 97 CALCIUM 1.
 FT METAL 102 114 CALCIUM 1.
 FT METAL 115 116 CALCIUM 1.
 FT METAL 119 122 CALCIUM 1.
 FT METAL 126 126 CALCIUM 1.
 FT METAL 131 142 CALCIUM 1.
 FT METAL 143 144 CALCIUM 1.
 FT METAL 146 150 CALCIUM 1.
 FT METAL 161 161 CALCIUM 1.
 FT METAL 162 164 CALCIUM 1.
 FT METAL 166 167 CALCIUM 1.
 FT METAL 176 177 CALCIUM 1.
 FT METAL 180 180 CALCIUM 1.
 FT METAL 182 183 CALCIUM 1.
 FT METAL 188 189 CALCIUM 1.
 FT METAL 192 195 CALCIUM 1.
 FT METAL 199 203 CALCIUM 1.
 FT METAL 204 206 CALCIUM 1.
 FT METAL 207 211 CALCIUM 1.
 FT METAL 214 231 CALCIUM 1.
 FT METAL 233 234 CALCIUM 1.
 FT METAL 237 246 CALCIUM 1.
 FT METAL 247 247 CALCIUM 1.
 FT METAL 249 249 CALCIUM 1.
 FT METAL 254 257 CALCIUM 1.
 FT METAL 258 259 CALCIUM 1.
 FT METAL 261 261 CALCIUM 1.
 FT METAL 264 267 CALCIUM 1.
 FT METAL 268 269 CALCIUM 1.
 SQ SEQUENCE 269 AA, 26623 MW, E8AF1A6A9A2676B CRC64;

Query Match 11.4%; Score 259; DB 1; Length 269;
 Best Local Similarity 31.5%; Pred. No. 4, 1e-10;
 Matches 96; Conservative 34; Mismatches 101; Indels 74; Gaps 14;

QY 13 AQNNYGLYGGQGVAVADTGLDTRGDSMHRAPRGRITALLY-ALGRTNNANDPNNGHTV 72
 DB 15 AAHNRGLTGSQVAVADTGLIST-----HPDLNTRGGASFPVGPB-STQDSNGHGHV 66

```
QY 73 AGSV--LGNALN-KOMAPQANLVFQSIMDSGGGLGGLPSNLNTLPSQAWNAGARIHTNSW 129
DB 67 AGTIALNNSITGLVGVAPSAELYAVKVLGADG--RGAISIAQGLEWAGNNGHVAHLSL 124
QY 130 GAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPGNSGTISAPGTAKNAITVGCATENYRP 189
DB 125 GSP---SPSATLEQAVNSATSRGVLVVAASGNSGASS--ISYPARYANAVAVGAT----- 174
QY 190 SFGSIADPNHIAQPSRGATDRGRIKPDVTAPGTFTLSARSLAPDSSFWANYSKYAY 249
DB 175 -----DONNRRASFQYAGL-----DIVAPGVNVQSTYP-----GSTYAS 210
QY 250 MGGTSMATPIVAGNVA-----OLREHFTKRGITPKPSLIKAAIAGATDVGL 297
DB 211 LNTGSMATPHVAGAAALVKQNPMSNVQIENH-LKN-----TATSLGSTNL-- 256
QY 298 GYPG 302
DB 257 -YGS 260

RESULT 13
THES_BACSP
ID THES_BACSP STANDARD; PRT; 401 AA.
AC Q45670;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Thermophilic serine proteinase precursor (EC 3.4.21.-) (Ak.1
DE protease).
OS Bacillus sp. (strain AK1).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1409;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95085262; PubMed=7993087;
RA Maciver B., McHale R.H., Saul D.J., Bergquist P.L.;
RT "Cloning and sequencing of a serine proteinase gene from a
RT thermophilic Bacillus species and its expression in Escherichia
RT coli.";
RL Appl. Environ. Microbiol. 60:3981-3988(1994).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.80 ANGSTROMS) OF 122-401.
RX MEDLINE=20057863; PubMed=1058904;
RA Smith C.A., Toogood H.S., Baker H.M., Daniel R.M., Baker E.N.;
RT "Calcium-mediated thermostability in the subtilisin superfamily: the
RT crystal structure of Bacillus Ak.1 protease at 1.8-A resolution.";
RL J. Mol. Biol. 294:1027-1040(1999).
CC -1- COFACTOR: Binds 3 calcium ions and 1 sodium ion per subunit.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: Has a pH optimum of 8.5, a temperature optimum of
CC 75 degrees Celsius.
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L29506; AAA63688.1; -.
CC PIR; I39974; I39974.
CC PDB; 1DBI; 18-NOV-99.
CC MEROPS; S08.009; -.
CC InterPro; IPR000209; Peptidase_S8.
CC InterPro; IPR009020; Protease_inhib.
CC Pfam; PF00082; Peptidase_S8; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC PROSITE; PS00136; SUBTILASE ASP; 1.
CC PROSITE; PS00137; SUBTILASE_HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; 1.
```

```
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Zymogen; Metal-binding; Calcium-binding;
FT SIGNAL; 3D-structure.
FT PROPEP 25 121
FT CHAIN 122 401
FT ACT_SITE 160 160
FT ACT_SITE 193 193
FT ACT_SITE 347 347
FT METAL 126 126
FT METAL 168 168
FT METAL 169 169
FT METAL 171 171
FT METAL 179 179
FT METAL 184 184
FT METAL 186 186
FT METAL 204 204
FT METAL 207 207
FT METAL 209 209
FT METAL 211 211
FT METAL 297 297
FT METAL 300 300
FT METAL 323 323
FT DISULFID 258 258
FT TURN 127 128
FT HELIX 129 132
FT TURN 135 136
FT HELIX 137 139
FT TURN 140 140
FT HELIX 141 144
FT TURN 145 147
FT TURN 152 153
FT STRAND 155 160
FT TURN 165 166
FT TURN 168 173
FT STRAND 174 179
FT TURN 180 183
FT STRAND 184 184
FT HELIX 193 202
FT STRAND 219 224
FT TURN 228 229
FT HELIX 234 246
FT TURN 247 248
FT STRAND 251 254
FT HELIX 263 274
FT TURN 275 276
FT STRAND 278 282
FT TURN 285 285
FT TURN 298 299
FT STRAND 301 306
FT TURN 308 309
FT STRAND 312 312
FT TURN 314 315
FT STRAND 316 316
FT TURN 320 321
FT STRAND 324 327
FT STRAND 331 335
FT TURN 336 338
FT STRAND 339 343
FT HELIX 346 362
FT TURN 363 364
FT HELIX 367 376
FT TURN 377 377
FT STRAND 379 379
FT TURN 382 383
FT STRAND 384 384
FT TURN 385 387
FT STRAND 388 388
FT TURN 391 392
FT HELIX 395 399
FT TURN 400 401
FT SEQUENCE 401 AA; 42835 MW; 1C736EFAA89F256F CRC64;
```


DR	PROSITE; PS00136;	SUBTILASE ASP; 1.
DR	PROSITE; PS00137;	SUBTILASE HIS; 1.
DR	PROSITE; PS00138;	SUBTILASE SER; 1.
KW	Hydrolase; Serine protease;	Cell wall; Peptidoglycan-anchor; Repeat;
KW	Signal.	
FT	SIGNAL	1 31
FT	CHAIN	32 1130
FT	PROPEP	1131 1167
FT	ACT_SITE	130 130
FT	ACT_SITE	193 193
FT	ACT_SITE	512 512
FT	DOMAIN	1029 1104
FT	DOMAIN	1034 1101
FT	REPEAT	1034 1050
FT	REPEAT	1051 1067
FT	REPEAT	1068 1084
FT	REPEAT	1085 1101
FT	SITE	1127 1131
FT	MOD_RES	1130 1130
SEQ	SEQUENCE	1167 AA; 128263 MW; D2DDC5E25752DA5D CRC64;
 Query Match 11.0%; Score 248.5; DB 1; Length 1167;		
Best Local Similarity 20.7%; Pred No 1.3e-08;		
Matches 127; Conservative 67; Mismatches 152; Indels 269; Gaps 25		
QY	21	GQGQVAVADTGLDTGRNDSMHEAFR-----GK 49
DB	121	GAGTVAVIDAGFDKN-----HEAWRLTDKTRARYQSKEDEKAKKEHGITYGEWVNDK 174
QY	50	ITALVALGRTNNANDPNGCHTHVAGSVLGNALN-----KGMAPOANLVFQSIMDSSG 101
DB	175	VAYXHDYSKDGKTAVDOEHGHVSGIUSGNAPSETKEPRLEGAEMPAQLLMRV-EIVN 233
QY	102	GLGLGFPSNLNTLFSAOWNAGARIHITNSGAPVNGAYT-AN----SRQVDYYVRNMDMTVL 156
DB	234	GLADYARNVAQAIRDVAIVLGAKVINMSFG--NAALAYANLPDETCKAFDYAKSGVSIV 290
QY	157	FAAGNEG-----PNSTGISAPTAKNAITVG-----182
DB	291	TSGAGNDSFGGKTRPLADHPDYGVWGTPAAADSTLTVASPDKOLTETAMVKTKDDQQD 350
QY	183	-----ATENVRPS-----FGSIA-----195
DB	351	KEMPVLSTRFEFPNKAYDIAYANGMKEDDFDKVKGIALLIERGDIDFKDKVANAKKAGA 410
QY	196	-----DNP-----NH 200
DB	411	VGVLIIYNQDKGFPIELFNVDQMPEAFTFSRKDGLLLKDNPKQTITFNATPKVLPTASGTK 470
QY	201	IAQFSRGATRDGRLKPDVTAPGTFFLSARSLLAPDSSFWANYSKYAVMGTSMATPIV 260
DB	471	LGRFFSWGLUTDGNLKPDIAPAQQDILL---SSVA-----NNKYAKLSGTSMSPALV 518
QY	261	AGNVAOLREHF-IKNRGITP-KFSLIKAAALIAGAT-----DVLGYPSGDQGWGRVTLID 312
DB	519	AGINGLLQOXYETQYVPDMPTEPSERLDLAKKVLMSSATALYDEDESKAFVSPRQOGAGAVDAK 578
QY	313	K-SLNVAYNEATALTTCQK-----ATYSFOTQAGKPLKI---SLVWTDAPGST 357
DB	579	KASAATMYVTDKD-NTSKSHLNNVSDKFVEVTVVHNSKDKQEELYQATVQTD-----631
QY	358	TASYTLVNDLDLVITAPNGQKYGVGNDFSYP-----YDNNWDGRNNVENVFINAPQSGTYTI 413
DB	632	-----KVDGGKHFAIPKVLVEASW-----QKITIPANSKKQVTV 665
QY	414	EVOQYNVPSGPORFS 428
DB	666	PIDA-----SRFS 673
 RESULT 15 WPRA_BACSU ID WPRA_BACSU STANDARD; PRZ; 894 AA.		

AC P54423; O06726;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cell wall-associated protease precursor (EC 3.4.21.-) [Contains: Cell
DE wall-associated polypeptides CWBP23 and CWBP52].
DE WPA OR BSU10770.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54 AND 414-428.
RC STRAIN=168;
RX MEDLINE=97158234; PubMed=9004506;
RA Margot P., Karamata D.;
RT "The wprA gene of Bacillus subtilis 168, expressed during exponential
RT growth, encodes a cell-wall-associated protease.";
RL Microbiology 142:3437-3444(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98015415; PubMed=9353931;
RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;
RT "Sequencing of regions downstream of addA (98 degrees) and citG (289
RT degrees) in Bacillus subtilis.";
RL Microbiology 143:3305-3308(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo R., Bortner M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.X., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guiseppi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinou S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroter R., Scoffone F., Soldo B.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat R., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -!- FUNCTION: NOT YET KNOWN; COULD BE INVOLVED IN PROTEOGLYCAN
CC DEGRADATION, BY CLEAVAGE OF ITS PEPTIDE BRIDGES.
CC -!- SUBCELLULAR LOCATION: Cell-wall bound.
CC -!- PTM: PROCESSED INTO CWBP23 AND CWBP52.
CC -!- SIMILARITY: Belongs to peptidase family S8.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
CC
DR EMBL; U58981; AAC25926.1; -
DR EMBL; Y09476; CAA70641.1; -
DR EMBL; Z99109; CAB12917.1; -
DR PIR; F69730; F69730.
DR HSRP; Q45670; 1DBI.
DR MROPS; S08.004; -
DR Subtilast; BG11846; wprA.
DR InterPro; IPR000209; Peptidase S8.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PRO0723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; FALSE_NEG.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Cell wall; Zymogen; Signal;
KW Complete proteome.
FT SIGNAL 1 31
FT CHAIN 32 894 CELL WALL-ASSOCIATED PROTEASE.
FT CHAIN 32 ? CWBP23.
FT PROPEP 2 413 POTENTIAL.
FT CHAIN 414 894 CWBP52.
FT ACT_SITE 466 466 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 650 650 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CONFLICT 9 9 V -> A (IN REF. 1).
FT CONFLICT 14 14 L -> I (IN REF. 1).
SQ SEQUENCE 894 AA; 96487 MW; 0F67C353B55F8DBC CRC64;

Query Match 10.9%; Score 246; DB 1; Length 894;
Best Local Similarity 23.6%; Pred. No. 1.4e-08;
Matches 103; Conservative 65; Mismatches 142; Indels 126; Gaps 17;
QY 25 VVAVADTGLDGTGRNDSNHEAFRGKITALYA---LGRTNNDPNHGHTHVAGSVLGNAL 81
DB 457 LIAVDTGVDSTLAD-----LKGVRTDLGHFVGRNNAMDDQGHGTHVAGIAAQSD 510
QY 82 N---KGMAPQANLVFQSMDSGGGLGFLPSNLTFLSQANNAGARIHTNSGAPVNGAY 137
DB 511 NGYSMTGLNAKAIIPVKVLDAG--SGDTQIALGKIYAADKGAIVNLSLG-----GGY 564
QY 138 TANSRQVDYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAITTVGATENYRPSFGSIADN 197
DB 565 SRVLEPALKYAADKXVLIATAASGNDGENA--LSYPASSKYTVSVGAT-----NR 611
QY 198 PNHIAQFSRGATRGRIKIPDVTAPGTFTLSARSLADSSFWANNYSKIYVNGTSHAT 257
DB 612 MDMTADFSNYGKGL-----DISAPGSDI-----PSLVP-----NGNVYMSGTSVAT 653
QY 258 PIVAGNVVAOLREHFIKNRGITPKPSLIKAALIAGATDVLGYPGSDQ----- 304
DB 654 PYAAAAGLL---FAQNPKL--KREVEDMLKKTADDISFESVDGGEELYDDYGDPIEI 708
QY 305 -----GMGRVTLDSLVNA---YVNSATALT-----GOKATY 334
DB 709 PKTPGVDMHSYGRLNVKAVSAADLQKVNKLESTQTAVRGSAGEKGTILIEVMNGKKKLG 768
QY 335 SPQTQAGKPLKSLVWTDAPGSTTASYTLVNDLDDVITAPNGSKYVGNDFSPYDNNWDG 394
DB 769 SAAGKDNKAFKNVIA-----TOKQOVLYLKATG-----DA 800
QY 395 RNNVNVFINAPQSGT 410
DB 801 KTSYKVVVWVGKPSGT 816

Search completed: March 31, 2004, 16:05:28
Job time : 8.49423 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 32.475 Seconds
(without alignments)
4206.909 Million cell updates/sec

Title: US-09-985-689A-3

Perfect score: 2263

Sequence: 1 NDVARGIVKADVAQNNGLY.....EVOAYNVSPQRFSLAIVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2263	100.0	433	2 Q9AQR4	Q9AQR4 bacillus sp
2	2251	99.5	433	2 Q9AQR1	Q9AQR1 bacillus sp
3	2234	98.7	433	2 Q9AQR2	Q9AQR2 bacillus sp
4	2020.5	89.3	434	2 Q9AQR0	Q9AQR0 bacillus sp
5	1998.5	88.3	639	2 Q9AQR3	Q9AQR3 bacillus sp
6	1994.5	88.1	640	2 Q9AQR9	Q9AQR9 bacillus sp
7	468	20.7	1825	5 Q9T9W1	Q9T9W1 dictyosteli
8	442.5	19.6	1702	5 Q9GTN7	Q9GTN7 dictyosteli
9	422.5	18.7	654	17 Q8UOC9	Q8UOC9 pyrococcus
10	399.5	17.7	561	16 Q8RBZ2	Q8RBZ2 thermoanaer
11	396.5	17.5	1239	16 Q8FBZ4	Q8FBZ4 streptomyce
12	377.5	16.7	430	16 Q8ENV1	Q8ENV1 oceanobacil
13	353	15.6	1253	16 Q9FCO6	Q9FCO6 streptomyce
14	347.5	15.4	412	2 Q9AER6	Q9AER6 thermoanaer
15	346.5	15.3	1237	2 Q8GGT4	Q8GGT4 streptomyce
16	345.5	15.3	412	16 Q8RC68	Q8RC68 thermoanaer

17	345.5	15.3	1102	2 P95694	P95694 streptomyce
18	344	15.2	444	16 Q9KB17	Q9KB17 bacillus ha
19	343.5	15.2	1105	2 Q8KH36	Q8KH36 streptomyce
20	336	14.8	1220	16 Q9L0A0	Q9L0A0 streptomyce
21	334	14.8	891	1 Q93635	Q93635 thermococcu
22	329	14.5	824	2 Q45464	Q45464 bacillus sp
23	329	14.5	1208	16 Q82B14	Q82B14 streptomyce
24	327	14.4	1139	16 Q82139	Q82139 streptomyce
25	327	14.4	1245	16 Q9RL54	Q9RL54 streptomyce
26	321	14.2	435	16 Q8EMJ3	Q8EMJ3 oceanobacil
27	315.5	13.9	442	16 Q31788	Q31788 bacillus su
28	307	13.6	1398	1 Q9P9L1	Q9P9L1 pyrococcus
29	304.5	13.5	799	16 Q9KEM1	Q9KEM1 bacillus ha
30	290.5	12.8	1135	1 Q9P9D1	Q9P9D1 uncultured
31	289.5	12.8	1407	16 Q816G4	Q816G4 bacillus ce
32	288	12.7	1345	1 Q54437	Q54437 staphylothe
33	284.5	12.6	959	16 Q8PMS7	Q8PMS7 xanthomonas
34	282.5	12.5	1101	16 Q82CF0	Q82CF0 streptomyce
35	279.5	12.3	431	2 Q9S3L6	Q9S3L6 bacillus sp
36	277.5	12.3	1098	16 Q9L1Z8	Q9L1Z8 streptomyce
37	272.5	12.0	1571	2 Q8GCW3	Q8GCW3 streptococc
38	269.5	11.9	715	2 P70765	P70765 alteromonas
39	269	11.9	434	2 Q54327	Q54327 bacillus sp
40	268	11.8	575	16 Q8ENW1	Q8ENW1 xanthomonas
41	266.5	11.8	1570	16 Q8E2V6	Q8E2V6 streptococc
42	266.5	11.8	1570	16 Q8DX06	Q8DX06 streptococc
43	266	11.8	530	2 Q8GBS2	Q8GBS2 vibrio sp.
44	266	11.8	586	16 Q8PAL8	Q8PAL8 xanthomonas
45	265	11.7	966	16 Q8PB28	Q8PB28 xanthomonas

ALIGNMENTS

RESULT 1
Q9AQR4 PRELIMINARY; PRT; 433 AA.
ID Q9AQR4
AC Q9AQR4;
DT 01-JUN-2001 (TREMREL. 17, Created)
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)
DT 01-JUN-2003 (TREMREL. 24, Last annotation update)
DE Protease (Fragment).
GN PROA.
OS Bacillus sp. D6.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=127889;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D6;
RX MEDLINE=20568675; PubMed=1118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus sp.: enzymatic properties, sequences, and
RT evolutionary relationships.";
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; AB046402; BAB21265.1; -.
DR HSSP; Q45670; 1DBI.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004283; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON TER 1
FT NON TER 433 433
SQ SEQUENCE 433 AA; 45636 MW; 52087E0A2516107F CRC64;

Query Match 100.0%; Score 2263; DB 2; Length 433;
 Best Local Similarity 100.0%; Pred. No. 2.4e-124;
 Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAAQNNYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAAQNNYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60

QY 61 NANDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 DB 61 NANDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120

QY 121 GARIHNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 DB 121 GARIHNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180

QY 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRKPDVTPAGTFFILSARSLAPDSSFW 240
 DB 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRKPDVTPAGTFFILSARSLAPDSSFW 240

QY 241 ANYNSKYAVMGCTSMATPIVAGNVAQLREHFKNRGITPKPSLIKAALIAGATDVGLGYP 300
 DB 241 ANYNSKYAVMGCTSMATPIVAGNVAQLREHFKNRGITPKPSLIKAALIAGATDVGLGYP 300

QY 301 SGDQGWGRVTLDKSLNVAVYNEATALTTGQKATYSFQAGKPLKISLVWTDAPGSTTAS 360
 DB 301 SGDQGWGRVTLDKSLNVAVYNEATALTTGQKATYSFQAGKPLKISLVWTDAPGSTTAS 360

QY 361 YTLVNDLDLVIITAPNGQKTVGNDPFSYPDNNWGRNNVENFINAPQSGTYTIEVQAYNV 420
 DB 361 YTLVNDLDLVIITAPNGQKTVGNDPFSYPDNNWGRNNVENFINAPQSGTYTIEVQAYNV 420

QY 421 PSGPQRFSLAIVH 433
 DB 421 PSGPQRFSLAIVH 433

RESULT 2
 Q9AQR1 PRELIMINARY; PRT; 433 AA.
 AC Q9AQR1
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROD.
 OS Bacillus sp. SD521.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=133780;
 RN [1]_TaxID=133780;
 RP SEQUENCE FROM N.A.
 RC STRAIN=SD521;
 RX MEDLINE=20568675; PubMed=11118284;
 RA Saeiki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 RA Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 RT alkaliphilic Bacillus spp.; enzymatic properties, sequences, and
 RT evolutionary relationships.";
 RT Biochem. Biophys. Res. Commun. 279:313-319(2000).
 CC -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046405; BAB21268.1; -.
 DR HSSP; Q45670; 1DBI.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.

QY 1 NDVARGIVKADVAAQNNYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAAQNNYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60

QY 61 NANDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
 DB 61 NANDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120

QY 121 GARIHNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
 DB 121 GARIHNSGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180

QY 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRKPDVTPAGTFFILSARSLAPDSSFW 240
 DB 181 VGATENYRPSFGSIADPNPHIAQFSSRGATRDGRKPDVTPAGTFFILSARSLAPDSSFW 240

QY 241 ANYNSKYAVMGCTSMATPIVAGNVAQLREHFKNRGITPKPSLIKAALIAGATDVGLGYP 300
 DB 241 ANYNSKYAVMGCTSMATPIVAGNVAQLREHFKNRGITPKPSLIKAALIAGATDVGLGYP 300

QY 301 SGDQGWGRVTLDKSLNVAVYNEATALTTGQKATYSFQAGKPLKISLVWTDAPGSTTAS 360
 DB 301 SGDQGWGRVTLDKSLNVAVYNEATALTTGQKATYSFQAGKPLKISLVWTDAPGSTTAS 360

QY 361 YTLVNDLDLVIITAPNGQKTVGNDPFSYPDNNWGRNNVENFINAPQSGTYTIEVQAYNV 420
 DB 361 YTLVNDLDLVIITAPNGQKTVGNDPFSYPDNNWGRNNVENFINAPQSGTYTIEVQAYNV 420

QY 421 PSGPQRFSLAIVH 433
 DB 421 PSGPQRFSLAIVH 433

RESULT 3
 Q9AQR2 PRELIMINARY; PRT; 433 AA.
 AC Q9AQR2
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Protease (Fragment).
 GN PROD.
 OS Bacillus sp. Y.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=133779;
 RN [1]_TaxID=133779;
 RP SEQUENCE FROM N.A.
 RC STRAIN=Y;
 RX MEDLINE=20568675; PubMed=11118284;
 RA Saeiki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
 RA Horikoshi K.;
 RT "Novel oxidatively stable subtilisin-like serine proteases from
 RT alkaliphilic Bacillus spp.; enzymatic properties, sequences, and
 RT evolutionary relationships.";
 RT Biochem. Biophys. Res. Commun. 279:313-319(2000).
 CC -/- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
 DR EMBL; AB046404; BAB21267.1; -.
 DR HSSP; Q45670; 1DBI.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF00082; Peptidase_S8; 1.

```
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Protease; Serine protease.
FT NON_TER 1
FT NON_TER 433
SQ SEQUENCE 433 AA; 45587 MW; B81291A803C775AE CRC64;

Query Match
Best Local Similarity 98.7%; Score 2234; DB 2; Length 433;
Matches 427; Conservative 3; Mismatches 12; Indels 0; Gaps 0;

QY 1 NDVARGIKADVQAQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITAYALGRTN 60
Db 1 NDVARGIKADVQAQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITAYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNALNKGWAPQANLVFQSIIMDSGGLGGLPSNLNTLFSQAWN 120
Db 61 NASDPNGHGHHTHVAGSVLGNALNKGWAPQANLVFQSIIMDSGGLGGLPSNLNTLFSQAWN 120
QY 121 GARIHTNSWGPVNGAYTANSRQVDVYVRNNDMTVLPFAAGNEGPNSTGISAPGTAKNAIT 180
Db 121 GARIHTNSWGPVNGAYTANSRQVDVYVRNNDMTVLPFAAGNEGPNSTGISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSIADNPNIHQAFSSRGATRDGRIKPDVTAPGTFFILSARSLAPDSF 240
Db 181 VGATENYRPSFGSIADNPNIHQAFSSRGATRDGRIKPDVTAPGTFFILSARSLAPDSF 240
QY 241 ANYNSKYAYMGTSNATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGY 300
Db 241 ANYNSKYAYMGTSNATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGY 300
QY 301 SGDQGWGRVTLDKSLNVAIVNEATALTGQKATYSFQAGKPKLSILVWTDAPGSTTA 360
Db 301 NGDQGWGRVTLDKSLNVAIVNEATALTGQKATYSFQAGKPKLSILVWTDAPGSTTA 360
QY 361 YTLVNDLDELVTAPNGQKVGNDPSPYDNNWGRNNVNFVINAPOSQGTIVIEQAYNV 420
Db 361 YTLVNDLDELVTAPNGQKVGNDPSPYDNNWGRNNVNFVINAPOSQGTIVIEQAYNV 420
QY 421 PSGPQRFSLAIHV 433
Db 421 PSGPQRFSLAIHV 433

RESULT 4
Q9AQR0
ID Q9AQR0 PRELIMINARY; PRT; 434 AA.
AC Q9AQR0;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Protease (Fragment).
GN PROE.
OS Bacillus sp. NV1.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=133781;
RN [1]_TaxID=133781;
RP SEQUENCE FROM N.A.
RC STRAIN=NV1;
RX MEDLINE=20568675; PubMed=11118284;
RA Saeki K., Okuda M., Hatada Y., Kobayashi T., Ito S., Takami H.,
RA Horikoshi K.;
RT "Novel oxidatively stable subtilisin-like serine proteases from
RT alkaliphilic Bacillus spp.: enzymatic properties, sequences, and
RT evolutionary relationships."
RL Biochem. Biophys. Res. Commun. 279:313-319(2000).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.
DR EMBL; AB046406; BAB21269.1; -.
DR HSSP; PC0782; 1SUP.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
```

```

DR HSSP; P00782; 1SUP.
DR GO; GO:0004289; F:subtilisin activity; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; P00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR SEQUENCE 639 AA; 68185 MW; 316AF5FFDE4EF54 CRC64;

Query Match      88.3%; Score 1998.5; DB 2; Length 639;
Best Local Similarity 88.0%; Pred. No. 1.1e-108;
Matches 382; Conservative 26; Mismatches 25; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNGYGLGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
Db 206 NDVARGIVKADVAQSSYGLGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 265
QY 61 NANDPNGHGHVAGSVLGN-ALNKGMAPQANLVFOSIMDSGGGLGGLPSNLTLFSQAWN 119
Db 266 NANDTNGHGHVAGSVLGN-ALNKGMAPQANLVFOSIMDSGGGLGGLPSNLTLFSQAWN 265
QY 120 AGARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTLVFAAGNEGPNSTISAPGTAKNAI 179
Db 326 AGARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTLVFAAGNEGPNSTISAPGTAKNAI 385
QY 180 TVGATENLRPSFGSADPNHIAQFSSRGATRDGRIKPDVTAPGTFFILSARSSLAPDSSF 239
Db 386 TVGATENLRPSFGSADPNHIAQFSSRGATRDGRIKPDVTAPGTFFILSARSSLAPDSSF 445
QY 240 WANYNSKYAYMGTSMATPIVAGNVAQLREHFVKRGITPKPSLIKAALIAGADVGLGY 299
Db 446 WANHDSKYAYMGTSMATPIVAGNVAQLREHFVKRGITPKPSLIKAALIAGADVGLGY 505
QY 300 PSQDQGWGRVTLDKSLNVAAYNEATLTGOKATYSFOTQAGKPLKISLWTDAPGSTTA 359
Db 506 PNGNGQGWGRVTLDKSLNVAAYNEATLTGOKATYSFOTQAGKPLKISLWTDAPGSTTA 565
QY 360 SVTLVNDLVLITAPNGQKYGNDPFSYPDNNWGRNNVNFVFNAPQSGTYTIEVQAYN 419
Db 566 SVTLVNDLVLITAPNGQKYGNDPFSYPDNNWGRNNVNFVFNAPQSGTYTIEVQAYN 625
QY 420 VPSGQRFSLAIVH 433
Db 626 VPGVQNFSLAIVN 639

RESULT 6
Q93UV9 PRELIMINARY; PRT; 640 AA.
AC Q93UV9;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Protease.
GN Protease.
OS Bacillus sp. KSM-KP43.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=109322;
RN [1] Itoh S.; Saeki K.;
RP SEQUENCE FROM N.A.
RC STRAIN=KP43;
RA Itoh S.; Saeki K.;
RT "new protease.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB051423; BAB55674.2;
DR GO; GO:0004289; F:subtilisin activity; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.

DR HSSP; P00782; 1SUP.
DR GO; GO:0004289; F:subtilisin activity; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; P00723; SUBTILISIN.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR SEQUENCE 640 AA; 67991 MW; 4BBAF77E9D592C15 CRC64;

Query Match      88.1%; Score 1994.5; DB 2; Length 640;
Best Local Similarity 87.8%; Pred. No. 1.9e-108;
Matches 380; Conservative 28; Mismatches 25; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNGYGLGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 60
Db 207 NDVARGIVKADVAQSSYGLGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRTN 266
QY 61 NANDPNGHGHVAGSVLGN-ALNKGMAPQANLVFOSIMDSGGGLGGLPSNLTLFSQAWN 119
Db 267 NANDTNGHGHVAGSVLGN-ALNKGMAPQANLVFOSIMDSGGGLGGLPSNLTLFSQAWN 326
QY 120 AGARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTLVFAAGNEGPNSTISAPGTAKNAI 179
Db 327 AGARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTLVFAAGNEGPNSTISAPGTAKNAI 386
QY 180 TVGATENLRPSFGSADPNHIAQFSSRGATRDGRIKPDVTAPGTFFILSARSSLAPDSSF 239
Db 387 TVGATENLRPSFGSADPNHIAQFSSRGATRDGRIKPDVTAPGTFFILSARSSLAPDSSF 446
QY 240 WANYNSKYAYMGTSMATPIVAGNVAQLREHFVKRGITPKPSLIKAALIAGADVGLGY 299
Db 447 WANHDSKYAYMGTSMATPIVAGNVAQLREHFVKRGITPKPSLIKAALIAGADVGLGY 506
QY 300 PSQDQGWGRVTLDKSLNVAAYNEATLTGOKATYSFOTQAGKPLKISLWTDAPGSTTA 359
Db 507 PNGNGQGWGRVTLDKSLNVAAYNEATLTGOKATYSFOTQAGKPLKISLWTDAPGSTTA 566
QY 360 SVTLVNDLVLITAPNGQKYGNDPFSYPDNNWGRNNVNFVFNAPQSGTYTIEVQAYN 419
Db 567 SVTLVNDLVLITAPNGQKYGNDPFSYPDNNWGRNNVNFVFNAPQSGTYTIEVQAYN 626
QY 420 VPSGQRFSLAIVH 433
Db 627 VPGVQNFSLAIVN 640

RESULT 7
Q8T9W1 PRELIMINARY; PRT; 1825 AA.
AC Q8T9W1;
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Serine protease/ABC transporter TagD.
GN TAGD.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ax4;
RA Anjard C.; Loomis W.F.;
RT "Evolution of the ABC transporters of Dictyostelium.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AF466309; AAL74253.1;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti.; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilisin activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001140; ABC_TM_transp.

```

```

DR InterPro; IPR003439; ABC transporter.
DR InterPro; IPR000209; Peptidase S8.
DR Pfam; PF00664; ABC membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PRODOM; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR KW ATP-binding; Protease; Transport.
SQ SEQUENCE 1825 AA; 202641 MW; E28160BC78613A3B CRC64;

Query Match      20.7%; Score 468; DB 5; Length 1825;
Best Local Similarity 27.1%; Pred. No. 1e-18;
Matches 157; Conservative 78; Mismatches 167; Indels 178; Gaps 23;

QY 19 LYGGQGVAVADTGLDGR-----NDS-----SMHEAFRGKITALYALGRTNNANDPNNGH 68
DB 19 LYGGQGVAVADTGLDGR-----NDS-----SMHEAFRGKITALYALGRTNNANDPNNGH 68
QY 327 LRGGQILSIADTGLDGHCHFFSDSNPIPNYNLNRKXVTVIGSL--HDNEDYVDGH 384
DB 327 LRGGQILSIADTGLDGHCHFFSDSNPIPNYNLNRKXVTVIGSL--HDNEDYVDGH 384
QY 69 GTHVAGSVLGN-----ALN--KGMAPQANLVFQSI--WDSSGGLGGLPSNLTLFSCAWN 119
DB 69 GTHVAGSVLGN-----ALN--KGMAPQANLVFQSI--WDSSGGLGGLPSNLTLFSCAWN 119
QY 385 GTHVCGSAGAPESSSLAISFSGLATDAKIAFPDLASDPNNPVPDPDYSQIQLFLYN 444
DB 385 GTHVCGSAGAPESSSLAISFSGLATDAKIAFPDLASDPNNPVPDPDYSQIQLFLYN 444
QY 120 AGARIHNSWGA-----PVNGAYTANSRQVDEYVRRN--DMTVLFAAGNEGPNSTGISAPGT 174
DB 120 AGARIHNSWGA-----PVNGAYTANSRQVDEYVRRN--DMTVLFAAGNEGPNSTGISAPGT 174
QY 445 AGARVHGDSWGLSGIQVGLGYSDDAGSIDDFLTHPFDLILRAAGNNEQYSSLLS-QAT 503
DB 445 AGARVHGDSWGLSGIQVGLGYSDDAGSIDDFLTHPFDLILRAAGNNEQYSSLLS-QAT 503
QY 175 AKNAITVGA----- 193
DB 175 AKNAITVGA----- 193
QY 504 AKNIVTGAETHTSYTTDALEYNPFETVAKSTLNSLCOSFDDDKYCTYTTACCTEYST 563
DB 504 AKNIVTGAETHTSYTTDALEYNPFETVAKSTLNSLCOSFDDDKYCTYTTACCTEYST 563
QY 184 -----TENYRPSFGSIAD-----NPNHIAOFFSRGATRDGRIKPDVTAPGTFILSA 229
DB 184 -----TENYRPSFGSIAD-----NPNHIAOFFSRGATRDGRIKPDVTAPGTFILSA 229
QY 564 VKGLSGCCTSVIKNSYASIRPSQPELYNENNICSFSKGPTHDGLKPDIVAFQVITSA 623
DB 564 VKGLSGCCTSVIKNSYASIRPSQPELYNENNICSFSKGPTHDGLKPDIVAFQVITSA 623
QY 230 RSSLA-----PDSFWANYSKYAYMGGTSMATPIVAGNVQALREH-----FIX 273
DB 230 RSSLA-----PDSFWANYSKYAYMGGTSMATPIVAGNVQALREH-----FIX 273
QY 624 RSGNANTDQCGDSL-PNTNALLS-ESGTSMATPLATAATTILRQYLVVDGYPTGSIVE 691
DB 624 RSGNANTDQCGDSL-PNTNALLS-ESGTSMATPLATAATTILRQYLVVDGYPTGSIVE 691
QY 274 NRGITPKPSLILKAALIAAGATDVGLGYPSGD-----GQWGRVTLDKS 314
DB 274 NRGITPKPSLILKAALIAAGATDVGLGYPSGD-----GQWGRVTLDKS 314
QY 682 SNKLOPTGSLKALMINNAQLNCTFLPSSNTNPNFNAVPTFAGANFVGWGLRMSSEW 741
DB 682 SNKLOPTGSLKALMINNAQLNCTFLPSSNTNPNFNAVPTFAGANFVGWGLRMSSEW 741
QY 315 LNVAIVNEA-----TALTGQKATYSF-----QTQA 340
DB 315 LNVAIVNEA-----TALTGQKATYSF-----QTQA 340
QY 742 L---YVSSGKPKPSRWVGIGELGDKKXASNWKKEYSLSTQNVSYCYFTYKPSSSGNSG 798
DB 742 L---YVSSGKPKPSRWVGIGELGDKKXASNWKKEYSLSTQNVSYCYFTYKPSSSGNSG 798
QY 341 GKP-LKISLVWTDAPGSTTASYTLVNDLVLVITAPNGQ-KYVGNDFSPYVDNN----- 391
DB 341 GKP-LKISLVWTDAPGSTTASYTLVNDLVLVITAPNGQ-KYVGNDFSPYVDNN----- 391
QY 799 GIPRIVATLVWTDPPSYGAKNLVNNLDLTMNTSESEFYFNSGGSSNGTKGTTLPL 858
DB 799 GIPRIVATLVWTDPPSYGAKNLVNNLDLTMNTSESEFYFNSGGSSNGTKGTTLPL 858
QY 392 WGRNNVENVF---INAPQSGTYTIEVQAVNPSPQRF 428
DB 392 WGRNNVENVF---INAPQSGTYTIEVQAVNPSPQRF 428
QY 859 QDSINNVEGIITPTNKSEISPRFIAGTNIPGPQNF 898
DB 859 QDSINNVEGIITPTNKSEISPRFIAGTNIPGPQNF 898

RESULT 8
Q9GTN7 PRELIMINARY; PRT; 1702 AA.
AC Q9GTN7;
DT 01-MAR-2001 (TremBrel. 16, Created)
DT 01-MAR-2001 (TremBrel. 16, Last sequence update)
DT 01-OCT-2003 (TremBrel. 25, Last annotation update)
DE TAGA.
GN TAGA.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RA Good J.R., Cabral M., Kuspa A.;
RT "taga", a putative serine protease/ABC transporter of Dictyostelium
RT that is expressed at the onset of development and is required for the
RT differentiation of a subpopulation of presore cells.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC 1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL; AF263455; AAG11416.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR000209; Peptidase S8.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PRODOM; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
DR KW ATP-binding; Transport.
SQ SEQUENCE 1702 AA; 187103 MW; 4A67716303CB7131 CRC64;

Query Match      19.6%; Score 442.5; DB 5; Length 1702;
Best Local Similarity 26.1%; Pred. No. 2.8e-17;
Matches 142; Conservative 91; Mismatches 168; Indels 143; Gaps 22;

QY 18 GLYGQGVAVADTGLDGR---NDSSMHEAFRGKITALYALGRTNNANDPNNGHGHVAG 74
DB 18 GLYGQGVAVADTGLDGR---NDSSMHEAFRGKITALYALGRTNNANDPNNGHGHVAG 74
QY 300 GIKGDEIVGCADTGIDINHCFFYDTNIGSTHRKLIIS-YSSGNGDQIDDEIDGHGTHVG 358
DB 300 GIKGDEIVGCADTGIDINHCFFYDTNIGSTHRKLIIS-YSSGNGDQIDDEIDGHGTHVG 358
QY 75 SVLGNAL-----NKGMAPQANLVFQSIIMSSGGLGGLPSNLTLFSCAWNAGARHTN 127
DB 75 SVLGNAL-----NKGMAPQANLVFQSIIMSSGGLGGLPSNLTLFSCAWNAGARHTN 127
QY 359 TIIGTTVDPPSVSEFSGGAPNSKVAFLVQVSGNGLSIQSNLTALYQSTYDQNAKHCD 418
DB 359 TIIGTTVDPPSVSEFSGGAPNSKVAFLVQVSGNGLSIQSNLTALYQSTYDQNAKHCD 418
QY 128 SMGAPVNGAYTANSRQVDEYVRRN--DMTVLFAAGNE---GPNS-GTISAPGTAKNAITVG 182
DB 128 SMGAPVNGAYTANSRQVDEYVRRN--DMTVLFAAGNE---GPNS-GTISAPGTAKNAITVG 182
QY 419 ANNSNIGFFYTGVTMIDRFQWDHPDFLVVRSGANNVFNFGNSIYTLTSSQESTSKNSLVVG 478
DB 419 ANNSNIGFFYTGVTMIDRFQWDHPDFLVVRSGANNVFNFGNSIYTLTSSQESTSKNSLVVG 478
QY 183 ATENYRPSFGSIAD----- 197
DB 183 ATENYRPSFGSIAD----- 197
QY 479 SSNQPSSSTYLSIDYWDWDFIYNISRTSVCTQGQSIYGITCSDVPTQTTSVDIQTQCCSN 538
DB 479 SSNQPSSSTYLSIDYWDWDFIYNISRTSVCTQGQSIYGITCSDVPTQTTSVDIQTQCCSN 538
QY 198 P-----NHIAQ-----FSSRGATRDGRIKPDVTAPGTFILSARSL 233
DB 198 P-----NHIAQ-----FSSRGATRDGRIKPDVTAPGTFILSARSL 233
QY 539 PILAKICSTEQOQYQNTSVYSEFIPSLFSGVGPTSDGLKPDLLAPGSPFISSR-SL 597
DB 539 PILAKICSTEQOQYQNTSVYSEFIPSLFSGVGPTSDGLKPDLLAPGSPFISSR-SL 597
QY 234 APDSFWANY-----NSKIAYMGGTSMATPIVAGNVQALRE-----HFINKR----- 275
DB 234 APDSFWANY-----NSKIAYMGGTSMATPIVAGNVQALRE-----HFINKR----- 275
QY 598 GPSSTI--NHCSPITSGIATGATAGAT--DVLGYPSGQWGRVTLDK-----SLNV-AYV 320
DB 598 GPSSTI--NHCSPITSGIATGATAGAT--DVLGYPSGQWGRVTLDK-----SLNV-AYV 320
QY 656 VGFQPSASLVKATLINTASINVDSTLEY---SQFGNQLSKLITTTNAQTSLDIPSSI 712
DB 656 VGFQPSASLVKATLINTASINVDSTLEY---SQFGNQLSKLITTTNAQTSLDIPSSI 712
QY 321 NEA-TALTGQKATYSFQTAGKPLKISLVWTDAPGSTTASYTLVNDLVLVITAPNGQKY 379
DB 321 NEA-TALTGQKATYSFQTAGKPLKISLVWTDAPGSTTASYTLVNDLVLVITAPNGQKY 379
QY 713 EKADPIINTGETNSYCFSLDSKADIDITLVWTDAPGSPSLSTFTLVNLDLALLA-----F 767
DB 713 EKADPIINTGETNSYCFSLDSKADIDITLVWTDAPGSPSLSTFTLVNLDLALLA-----F 767
QY 380 VGNDFSPYVDNN-----WDGRNNVENVFINAPQSGTYTIEVQAVNPSPQRF 427
DB 380 VGNDFSPYVDNN-----WDGRNNVENVFINAPQSGTYTIEVQAVNPSPQRF 427
QY 768 VDGELSI-YSGNSETFIKNTSQVIFPQLNNVIRIKDAP-GSYDVKIFGTNIPNQSY 826
DB 768 VDGELSI-YSGNSETFIKNTSQVIFPQLNNVIRIKDAP-GSYDVKIFGTNIPNQSY 826
QY 428 SLAI 431
DB 428 SLAI 431
QY 827 SVVI 830
DB 827 SVVI 830

```

```
RESULT 9
Q8UOC9 PRELIMINARY; PRT; 654 AA.
AC Q8UOC9;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Alkaline serine protease.
GN PF1670.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;
RN [1]_TaxID=2261;
RP SEQUENCE FROM N.A.
RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE010265; AAL61794.1; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase S8.
DR Pfam; PF00082; Peptidase S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Protease, Complete proteome.
SQ SEQUENCE 654 AA; 70230 MW; 1CB145A5F505DB34 CRC64;

Query Match 18.7%; Score 422.5; DB 17; Length 654;
Best Local Similarity 31.4%; Pred. No. 1.2e-16;
Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;

QY 16 NYGLYGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRNNAN-----DPNGHG 69
DB 152 NLGYDGSIGITIDTGD-----ASHPDLQGV-----IGWVDFVNGRSYPYDDRHG 200
QY 70 THVAGSVLG-----NALNKGMAPOANLVFQSIM--DSGGGLGLPLSNLNTLFSQAWNAGA 122
DB 201 TEVASIAGTGAASCKYKGNAPGAKLAGIKVLGADGSGSISTIKGVAVENKDKYGI 260
QY 123 RIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPOTAKNAITVG 182
DB 261 KVINLSLGSQSSDGTDLALSQAVNAADAGLVVVVAGNSGPNKYTTIGSPAAASKVITVG 320
QY 183 ATENYRPSFGSIADPNHIAOFSSRGATDGRIPKDYTPAGTFTILSARSLAPDSFWAN 242
DB 321 AVDKY-----DVTSSRGPTADGRLPKPEVVAPGNWIIAARAS---GTSWGP 366
QY 243 YNSKYAYNGTSMATPIVAGNVAQLREHFIRNGITP---KPSLIKAALIAGATDVG-LG 298
DB 367 INDYTYAAPGTSMATPHVAGIALLQ---AHPSTWPKVKXTALLETADIVKPEDEIADIA 423
QY 299 YPSGQDQGRVTLDKSLNVAYNENATALTGQKA-----TYSFQTQAGKPLKISLWTD 353
DB 424 Y-----GAGRNVAYKAIN--VDNYAKLVFTGVANKSQTHQFVISGASFYATLYWDNA 476
QY 354 FGSSTASYTLVNDLVLITAPNGQRYGVNDPFSYPYDNNWDRNENNVFINAPQSGTYTI 413
DB 477 N-----SDLDLYLDPNQV---DYSY-----TAYGFEKVGYNPTDGTWTI 517
QY 414 EVQAVNVPSGQRFSLAIV 432
DB 518 KTVSY---SGSANYQDVV 533
```

```
RESULT 10
Q8RBJ2
```

```
Q8RBJ2 PRELIMINARY; PRT; 561 AA.
AC Q8RBJ2;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Subtilisin-like serine proteases.
GN APR22 OR TTE0824.
OS Thermoanaerobacter tengcongensis.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=119072;
RN [1]_TaxID=119072;
RP SEQUENCE FROM N.A.
RC STRAIN=MB4 / JCM 11007;
RX MEDLINE=21992816; PubMed=11997336;
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
Tan H., Chen R., Wang J., Yu J., Yang H.;
RT "A complete sequence of T. tengcongensis genome.";
RL Genome Res. 12:689-700(2002).
DR EMBL; AE0113049; AAM24081.1; -.
DR GO; GO:0005503; F:calcium ion binding; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR002029; Peptidase_S8.
DR InterPro; IPR007280; PPC.
DR InterPro; IPR009020; Protease_inhib.
DR Pfam; PF00082; Peptidase_S8; 1.
DR Pfam; PF04151; PPC; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Protease, Complete proteome.
SQ SEQUENCE 561 AA; 59968 MW; BA9C5C52F7083A18 CRC64;

Query Match 17.7%; Score 399.5; DB 16; Length 561;
Best Local Similarity 30.8%; Pred. No. 2.1e-15;
Matches 141; Conservative 64; Mismatches 160; Indels 93; Gaps 19;

QY 6 GIVKADVAQNNYGLYGQGVAVADTGLDGRNDSMHEAFRGKITALYALGRNNAN-- 63
DB 155 GITK---AKSDFGVTKMTITAITDIDGNHVDLS-----GGKI-----IGWKDFINNK 201
QY 64 ----DPNGHGTHV-----AGSVLGNALNKGMAPOANLVFQSIMDSSG-----GLGGL 106
DB 202 TTPYDDNGHGTHVASIAAGTGAGNSFYKGVAPDALLVGIKVLGDANGSGSMSTVTAGIDWA 261
QY 107 PSNLTNLFSAWNAGARIHTNSWAGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGNS 166
DB 262 VQNKDVGKVINLSLGTSTSSDG--TDSTSLAVNRAVD-----SGIVVVVAAAGNSGPAK 314
QY 167 GTISAPGTAKNAITVGATENYRPSFGSIADPNHIAOFSSRGATDGRIPKDYTPAGTFTI 226
DB 315 YTIISGFAEKAITVAAD----VGLGFN---LASFSSRGPTADGRIPKDIAPGNYI 367
QY 227 LGAARSLAPDSFWANYSKYAYMGGTSMATPIVAGNVAQLREHFIRNGITPESLIKA 286
DB 368 TAAK-----ANSVNGYVYTSGTSMATPIVAGTVALMLN---ANPNLTNDPA---KN 412
QY 287 ALIAGATDVLGYPGSDQGRVTLDKSLNVA-----YVNEATALTGQKAT 333
DB 413 IIMSTAKSGPPSKVDYAGRLDGYEARVAGNFRGNNDVPNHYISGLVPSGRYSDT 472
QY 334 YSFQ-QQAGKPLKISLWTDAPGSTASYTLVNDLVLITAPNGQRYGVNDPFSYPYDNNW 392
DB 473 WTFNATNYSYPIAITLIIPD-----WANYN--PDFIYLYDPSPGTLIKSS----- 515
QY 393 DGRNENNVFINAPQSGTYTIEVQAVNVPSGQRFSLA 430
```


516 TGTQRTTILPSTGTGYVVKVSYR-GSGNYFFDL 552

Db 219 VEADLADSTAGIAGAPRAWAGNTGGQVEVAVLDTGVDAG-----HPDLADRIARQSFV 272

QY 58 RTNNANDPNHGHTHVAGSVLGNAL-----NKGMAPOANLVFQSIMDSSGGLGLPNSLNT 112

Db 273 PDENTDDRDGHGHTHVASTIAGTGAASAGKEGVAPCARLSIGKVLDS-GRGQISWTLAA 331

QY 113 LFSOAWNAGARIHTNSWGA-PVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISA 171

Db 332 MEWAVERHAKIVNMSLGSGSDGSDPMSRAVDRLSATGALFVVAAGN-GGEAGSIGA 390

QY 172 PGTAKNAITVGTATENYRPFSGSIADPNHIAQFSSRGATRDGRKIDKPDVAPGTFTILSARS 231

Db 391 PGVATSALTGVA-----VDATDTLAPSSQGPVVDGALKPEITAPGVGILAA-- 437

QY 232 SLAPDSSFWMYNYKIAMVGTSMATPIVAGNVAQREHFKNRGITPTPSPILKAL----- 288

Db 438 ----NSSFAAGNGAYQSLGTSMTATPHVAGAAALL-----AAAPEDLSGSGALKDV 484

QY 289 IAGATDVGLGSPGDCQGWGRVTLDKSLNVAYNNEATLTGQK-----ATYSFOTQAG 341

Db 485 LASSGHRTPRYDAFOAGSGRVDVDAVRAVGYASATAYAPGSGPGRVRLVTVTTTGA 544

QY 342 KPLKISLWTDAPGST---TASYTLV---NDLVLVT-----APNGQKYVGNDFSTPYDN 390

Db 545 VTLELSVAATHAPEGVEFLSASRTVPAHGTADVTLLTDGSGAGRAYSGQILA---T 600

QY 391 NWDGRNNVENVFINAPQSGTYTIEVQAVNVVPSGPQRFSLAI 431

Db 601 DADARN-----VAHTAVSAGPVRHKUTV 623

RESULT 12

Q8ENV1 PRELIMINARY; PRT; 430 AA.

AC Q8ENV1

DT 01-MAR-2003 (TremBLrel. 23, Created)

DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)

DE Intracellular alkaline serine proteinase.

GN OB2375.

OS Oceanobacillus iheyensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.

OX NCBI_TaxID=182710;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HTS831 / DSM 14371 / JCM 11309;

RX MEDLINE=22220767; PubMed=12235376;

RA Takami H., Takaki Y., Uchiyama I.;

RT Genome sequence of Oceanobacillus iheyensis isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments."

RT Nucleic Acids Res. 30:3927-3935(2002).

RL EMBL; AP004601; BAC14331.1; "

DR GO; GO:0004289; P:subtilase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR000209; Peptidase_S8.

DR Pfam; PF00082; Peptidase_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE ASP; 1.

DR PROSITE; PS00137; SUBTILASE HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.

KW Complete proteome.

SQ SEQUENCE 430 AA; 45938 MW; 6D09A99B8C1E110F CRC64;

Query Match 16.7%; Score 377.5; DB 16; Length 430;

Best Local Similarity 34.5%; Pred. No. 2.8e-14;

Matches 119; Conservative 48; Mismatches 119; Indels 59; Gaps 16;

QY 2 DVARGIVKADVAQNNYGLYGQGVVAADTGLDTRNDSSMHEAFRGKIT--ALYALGRT 59

Db 121 DTASSINADLVKES-GLTGGGSTIAVIDTGHF-----HBDLEGRIGFADFVKGT 172

QY 60 NNANDPNHGHTHVAGSVLGN-ALN-----KGMAPQANLVFQSIMDSSGGLGLPNS- 108

516 TGTQRTTILPSTGTGYVVKVSYR-GSGNYFFDL 552

Db 219 VEADLADSTAGIAGAPRAWAGNTGGQVEVAVLDTGVDAG-----HPDLADRIARQSFV 272

QY 58 RTNNANDPNHGHTHVAGSVLGNAL-----NKGMAPOANLVFQSIMDSSGGLGLPNSLNT 112

Db 273 PDENTDDRDGHGHTHVASTIAGTGAASAGKEGVAPCARLSIGKVLDS-GRGQISWTLAA 331

QY 113 LFSOAWNAGARIHTNSWGA-PVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISA 171

Db 332 MEWAVERHAKIVNMSLGSGSDGSDPMSRAVDRLSATGALFVVAAGN-GGEAGSIGA 390

QY 172 PGTAKNAITVGTATENYRPFSGSIADPNHIAQFSSRGATRDGRKIDKPDVAPGTFTILSARS 231

Db 391 PGVATSALTGVA-----VDATDTLAPSSQGPVVDGALKPEITAPGVGILAA-- 437

QY 232 SLAPDSSFWMYNYKIAMVGTSMATPIVAGNVAQREHFKNRGITPTPSPILKAL----- 288

Db 438 ----NSSFAAGNGAYQSLGTSMTATPHVAGAAALL-----AAAPEDLSGSGALKDV 484

QY 289 IAGATDVGLGSPGDCQGWGRVTLDKSLNVAYNNEATLTGQK-----ATYSFOTQAG 341

Db 485 LASSGHRTPRYDAFOAGSGRVDVDAVRAVGYASATAYAPGSGPGRVRLVTVTTTGA 544

QY 342 KPLKISLWTDAPGST---TASYTLV---NDLVLVT-----APNGQKYVGNDFSTPYDN 390

Db 545 VTLELSVAATHAPEGVEFLSASRTVPAHGTADVTLLTDGSGAGRAYSGQILA---T 600

QY 391 NWDGRNNVENVFINAPQSGTYTIEVQAVNVVPSGPQRFSLAI 431

Db 601 DADARN-----VAHTAVSAGPVRHKUTV 623

RESULT 12

Q8ENV1 PRELIMINARY; PRT; 430 AA.

AC Q8ENV1

DT 01-MAR-2003 (TremBLrel. 23, Created)

DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)

DE Intracellular alkaline serine proteinase.

GN OB2375.

OS Oceanobacillus iheyensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.

OX NCBI_TaxID=182710;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HTS831 / DSM 14371 / JCM 11309;

RX MEDLINE=22220767; PubMed=12235376;

RA Takami H., Takaki Y., Uchiyama I.;

RT Genome sequence of Oceanobacillus iheyensis isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments."

RT Nucleic Acids Res. 30:3927-3935(2002).

RL EMBL; AP004601; BAC14331.1; "

DR GO; GO:0004289; P:subtilase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR000209; Peptidase_S8.

DR Pfam; PF00082; Peptidase_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE ASP; 1.

DR PROSITE; PS00137; SUBTILASE HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.

KW Complete proteome.

SQ SEQUENCE 430 AA; 45938 MW; 6D09A99B8C1E110F CRC64;

Query Match 17.5%; Score 396.5; DB 16; Length 1239;

Best Local Similarity 30.8%; Pred. No. 9e-15;

Matches 141; Conservative 55; Mismatches 172; Indels 93; Gaps 16;

QY 8 VKADVAQNNY-----GLYGQGVVAADTGLDTRNDSSMHEAFRGKITALYALG 57

516 TGTQRTTILPSTGTGYVVKVSYR-GSGNYFFDL 552

Db 219 VEADLADSTAGIAGAPRAWAGNTGGQVEVAVLDTGVDAG-----HPDLADRIARQSFV 272

QY 58 RTNNANDPNHGHTHVAGSVLGNAL-----NKGMAPOANLVFQSIMDSSGGLGLPNSLNT 112

Db 273 PDENTDDRDGHGHTHVASTIAGTGAASAGKEGVAPCARLSIGKVLDS-GRGQISWTLAA 331

QY 113 LFSOAWNAGARIHTNSWGA-PVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISA 171

Db 332 MEWAVERHAKIVNMSLGSGSDGSDPMSRAVDRLSATGALFVVAAGN-GGEAGSIGA 390

QY 172 PGTAKNAITVGTATENYRPFSGSIADPNHIAQFSSRGATRDGRKIDKPDVAPGTFTILSARS 231

Db 391 PGVATSALTGVA-----VDATDTLAPSSQGPVVDGALKPEITAPGVGILAA-- 437

QY 232 SLAPDSSFWMYNYKIAMVGTSMATPIVAGNVAQREHFKNRGITPTPSPILKAL----- 288

Db 438 ----NSSFAAGNGAYQSLGTSMTATPHVAGAAALL-----AAAPEDLSGSGALKDV 484

QY 289 IAGATDVGLGSPGDCQGWGRVTLDKSLNVAYNNEATLTGQK-----ATYSFOTQAG 341

Db 485 LASSGHRTPRYDAFOAGSGRVDVDAVRAVGYASATAYAPGSGPGRVRLVTVTTTGA 544

QY 342 KPLKISLWTDAPGST---TASYTLV---NDLVLVT-----APNGQKYVGNDFSTPYDN 390

Db 545 VTLELSVAATHAPEGVEFLSASRTVPAHGTADVTLLTDGSGAGRAYSGQILA---T 600

QY 391 NWDGRNNVENVFINAPQSGTYTIEVQAVNVVPSGPQRFSLAI 431

Db 601 DADARN-----VAHTAVSAGPVRHKUTV 623

RESULT 12

Q8ENV1 PRELIMINARY; PRT; 430 AA.

AC Q8ENV1

DT 01-MAR-2003 (TremBLrel. 23, Created)

DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)

DE Intracellular alkaline serine proteinase.

GN OB2375.

OS Oceanobacillus iheyensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.

OX NCBI_TaxID=182710;

RN [1]

RP SEQUENCE FROM N

516 TGTQRTTILPSTGTGYVVKVSYR-GSGNYFFDL 552

Db 219 VEADLADSTAGIAGAPRAWAGNTGGQVEVAVLDTGVDAG-----HPDLADRIARQSFV 272

QY 58 RTNNANDPNHGTHVAGSVLGNAL-----NKGMAPOANLVFQSIMDSSGGLGLPNSLNT 112

Db 273 PDENTDDRDGHGTHVASTIAGTGAASAGKEGVAPCARLSIGKVLDSN-GRGQISWTLAA 331

QY 113 LFSOAWNAGARIHTNSWGA-PVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISA 171

Db 332 MEWAVERHAKIVNMSLGSGSDGSDPMSRAVDRLSATGALFVVAAGN-GGEAGSIGA 390

QY 172 PGTAKNAITVGATENYRPFSGSIADPNHIAQFSSRGATRDGRKIDKPDVTPAGTFTILSARS 231

Db 391 PGVATSALTGVA-----VDATDTLAPSSQGPVVDGALKEITAPGVGILAA-- 437

QY 232 SLAPDSSFWMYNYKIAMGTSMAPIVAGNVAQREHFKNRGITPTPSPILKAL----- 288

Db 438 ----NSSFAAGNGAYQSLGTSMTATPHVAGAAALL-----AAAPEDLSGSKADV 484

QY 289 IAGATDVGLGSPGDCQGWGRVTLDKSLNVAYNNEATLTGQK-----ATYSFOTQAG 341

Db 485 LASSGHRTPRYDAFOAGSGRVDVDAVRAVGYASATAYAPGSGPVRRLVTVTTTGA 544

QY 342 KPLKISLWTDAPGST---TASYTLV---NDLVLVT-----APNGQKYVGNDFSTPYDN 390

Db 545 VTLELSVAATHAPEGVEFLSASRTVPAHGTADVTLLTDGSGAGRAYSGQILA---T 600

QY 391 NWDGRNNVENVFINAPQSGTYTIEVQAVNVPSGPQRFSLAI 431

Db 601 DADARN-----VAHTAVSAGPVRHKUTV 623

RESULT 12

Q8ENV1 PRELIMINARY; PRT; 430 AA.

AC Q8ENV1

DT 01-MAR-2003 (TremBLrel. 23, Created)

DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)

DE Intracellular alkaline serine proteinase.

GN OB2375.

OS Oceanobacillus iheyensis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.

OX NCBI_TaxID=182710;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HTS831 / DSM 14371 / JCM 11309;

RX MEDLINE=22220767; PubMed=12235376;

RA Takami H., Takaki Y., Uchiyama I.;

RT Genome sequence of Oceanobacillus iheyensis isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments."

RT Nucleic Acids Res. 30:3927-3935(2002).

RL EMBL; AP004601; BAC14331.1; "

DR GO; GO:0004289; P:subtilase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR000209; Peptidase_S8.

DR Pfam; PF00082; Peptidase_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE ASP; 1.

DR PROSITE; PS00137; SUBTILASE HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.

KW Complete proteome.

SQ SEQUENCE 430 AA; 45938 MW; 6D09A99B8C1E110F CRC64;

Query Match 16.7%; Score 377.5; DB 16; Length 430;

Best Local Similarity 34.5%; Pred. No. 2.8e-14;

Matches 119; Conservative 48; Mismatches 119; Indels 59; Gaps 16;

QY 2 DVARGIVKADYQNNYGLYGQGVVAADTGLDTRNDSSMHEAFRGKIT--ALYALGRT 59

Db 121 DTASSINADLVKES-GLTGGGSIIVDTGHP-----HEDLEGRIGFADFVKGT 172

QY 60 NNANDPNHGTHVAGSVLGN-ALN-----KGMAPQANLVFQSIMDSSGGLGLPNS- 108

Db 516 TGTQRTTILPSTGTGYVVKVSYR-GSGNYFFDL 552

RESULT 11

Q9FBZ4 PRELIMINARY; PRT; 1239 AA.

AC Q9FBZ4

DT 01-MAR-2001 (TremBLrel. 16, Created)

DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)

DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)

DE Putative secreted peptidase.

DE SC07188 OR SC9A11.16C.

GN Streptomyces coelicolor.

OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Saunders D.C., Harris D.;

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,

RA Kinashi H., Hopwood D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2) / M145;

RC MEDLINE=97000951; PubMed=8843436;

RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,

RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,

RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2).";

RL Nature 417:141-147(2002).

RL EMBL; AL939130; CAC01588.1; "

DR HSP; Q99405; IMPT.

DR GO; GO:0008233; P:peptidase activity; IEA.

DR GO; GO:0004289; P:subtilase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR03137; PA.

DR InterPro; IPR000209; Peptidase_S8.

DR Pfam; PF02225; PA; 1.

DR Pfam; PF00082; Peptidase_S8; 1.

DR PRINTS; PR00723; SUBTILISIN.

DR PROSITE; PS00136; SUBTILASE ASP; 1.

DR PROSITE; PS00137; SUBTILASE HIS; 1.

DR PROSITE; PS00138; SUBTILASE_SER; 1.

KW Complete proteome.

SQ SEQUENCE 1239 AA; 128505 MW; 8F5E9AC68EB1260A CRC64;

Query Match 17.5%; Score 396.5; DB 16; Length 1239;

Best Local Similarity 30.8%; Pred. No. 9e-15;

Matches 141; Conservative 55; Mismatches 172; Indels 93; Gaps 16;

QY 8 VKADVAQNNY-----GLYGQGVVAADTGLDTRNDSSMHEAFRGKITALYALG 57

Db 173 EYDD-NHGTTCAGDAAGSAGQYQGPAPDANLVGVKNTGS-GSLSTVIEGID 230
 QY 109 -----NLNLTFSQAWNAGARHTNSWGAPVNGAYTANSRQVDEYVRNNDMTLFAA 159
 Db 231 WCIONQSKYNIL--SLSLGSDATEPAEGDPVNAV-----ETAWNGVVCVAA 279
 QY 160 NNEGNSGTSISAPGTAKNAITVGATENVRPSPGSTADPNHIAQFSSRGATRGRIKPDV 219
 Db 280 GNSGPGDKTVGSPGSPKVTITGAADDNNTAERS-----DDSVAFSSRGPTIDGLTKFNL 335
 QY 220 TAPGFFILSARS--SLAPDSSFWANYSKYAMGSGTSMATPIVAGNVAQLREHFIKNGI 277
 Db 336 LPEFGDIVSLRAPGSFIDKTNKSARVGSYISLSTGTSNATPICAGIVAQLLQ---SDSSL 392
 QY 278 TPKPSLIIKAAIAGATDVLGHPGSDQGWGRVTLKSLNVAVNE 322
 Db 393 T--PNQVKELMEACQD--LQSPNVQAGYL---NAANLINE 430

RESULT 13
 Q9FC06 PRELIMINARY; PRT; 1253 AA.
 AC Q9FC06
 DT 01-MAR-2001 (TremBLrel. 16, Created)
 DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE Putative secreted peptidase.
 GN SCO7176 OR SC8A11.04C.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Saunders D.C., Harris D.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8943436;
 RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RL Mol. Microbiol. 21:77-96 (1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdeno-Narraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)."
 RL Nature 417:141-147 (2002).
 RL EMBL; AL939130; CAC01576.1; --
 DR HSSP; Q99405; IMPT.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0004289; F:subtilase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR003137; PA.
 DR InterPro; IPR000209; Peptidase_S8.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF00082; Peptidase_S8; 1.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; 1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 KW Complete proteome.
 SQ SEQUENCE 1253 AA; 130971 MW; AA69B417EFEEDB89 CRC64;
 Query Match 15.6%; Score 353; DB 16; Length 1253;
 Best Local Similarity 30.1%; Pred. NO. 3, 1e-12;
 Matches 143; Conservative 56; Mismatches 164; Indels 112; Gaps 25;
 QY 18 GYGQGVVAVADTGLDGTGRNDSMHEAFRGKITAYALGRNTNANDPNHGHTHVAGSVL 77
 Db 235 GNTGEGVAVLDTGVDAG-----HPDPAGRIATAAGFVDPDQVTDNRNGHGHVASTVA 288
 QY 78 G-----NALNKGMAPQANLVFQSIMDSSGGLGLSNLNTLFSQAW---NAGARIHTNSW 129
 Db 289 GTGAASGGVEKGVAPGASLHIGKVLDSG--SGQSWV--LAGMEWAVRDQAKIVMSL 344
 QY 130 G-APVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAITVGATENYR 188
 Db 345 GDSPTDGT-DPLSEAVNMLSAETGALFVVAAGNSGPEAVTGTTPAADAAALTVGAVNG-- 401
 QY 189 PSFGSIADPNHIAQFSSRG-ATRDGRKPKDVTAGTILSNARSSLPADSSFWANYSKY 247
 Db 402 PKGK-----VDQLADFFSSRGPRVGNVAKPDLTAPGVGLAARSYAPEG-----EGAY 450
 QY 248 AYMGGTSMATPIVAGNVAQL-REHFIKNGITPKP-----SLIKAAIAGATDVLGYPGSG 302
 Db 451 QSLSGTSMATPHVAGAAALLAAEH-----PDWTGQRLKEALV-GTTAGTQRFSPF 499
 QY 303 DOGWGRVTLKSLNVAVNEATALTQO---KATVSF-----QTQAG-KPLKIS 347
 Db 500 DAGSGRV-----DVAAAVRSTLLASGDAFAQAHPYTPGQTRRDVDTYNSGPAFALD 553
 QY 348 LVMTDA-----PGSTTASYTLVNDLD-----LVITAPNGQKYV 380
 Db 554 LALSFAELPEGLFTLSEAQTVPAGHTASVGVITHLDAEDNGAVATRLVASGADGAVIA 613
 QY 381 GNDFSYPYDNNWDGNNVENVFINA-----PQSGTYTIEVQAVNVPSPGPFSL 429
 Db 614 RT-----PVGWKEGRR--ATLALTAKHDKHDKPLSGTILKDVERN--TAPKYYSV 660

RESULT 14
 Q9AER6 PRELIMINARY; PRT; 412 AA.
 AC Q9AER6
 DT 01-JUN-2001 (TremBLrel. 17, Created)
 DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
 DE Thermicidin (Subtilisin-like protease thermicidin).
 OS Thermoanaerobacter yonsei.
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.
 OX NCBI_TaxID=111519;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jiang H.J., Kim B.C., Pyun Y.R., Kim Y.S.;
 RT "A novel subtilisin-like serine protease from Thermoanaerobacter
 yonseiensis K8-1: cloning, expression and biochemical properties."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Jiang H.J., Kim B.C., Pyun Y.R., Kim Y.S.;
 RT "Subtilisin-like protease, thermicidin, from Thermoanaerobacter
 yonsei." (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.

```

DR EMBL; AY028704; AAK27733.1; -.
DR EMBL; AF305633; AAL09366.1; -.
DR HSP; Q45670; 1DBI.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PROSITE; PS00136; SUBTILASE ASP; 1.
DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
DR Hydrolase; Protease; Serine protease.
KW SEQUENCE 412 AA; 44503 MW; E3C6A0F081B1A1D47 CRC64;

Query Match      15.4%; Score 347.5; DB 2; Length 412;
Best Local Similarity 35.9%; Pred. No. 1.5e-12;
Matches 117; Conservative 39; Mismatches 113; Indels 57; Gaps 17;

QY 15 NNYGLYGQGVVAVADTGLDTRNDSSMHEAF---RQKITALY-ALGRTNNANDPNGHGT 70
DB 116 NDLGYTGKGTIAFLDTGI-----YHPHDFTRPKRIIAFYDVVNGKKQPYDDNGHGT 168
QY 71 HVAGSVLCN--ALN--KGMAPQANLVFQSIMDSSGGLGLPSNLNTLFSQAW-----N 119
DB 169 HVAGDAAGNYSNGKYKGVAPENIVAVKLDY-----GRGSSDILAGMQWVLDNKEK 224
QY 120 AGARIHTNSWG-APVNGAYTANSQVDEYVR-----NNDMTVLFAGNENGPNGTISAP 172
DB 225 YNIRIVLSISIGETPALPTF-----LDPLVRGVDTLWKGIVVYVVAAGNSGPNYSITSP 278
QY 173 GTAKNAITVGATENYRPSFGSIADPNPHIAQFSSRGATROGRIKPDVTAPGRTILSARSS 232
DB 279 GTSRNAITVGAVDKRTP--DIED--DEVAKFSRGGPY--LYKPDVVAPEVKLVSTAG 332
QY 233 LAPSSFWANNYSKYANGGTSMTPIVAGNVAQLREHFIKNRGITPKPSLIXAALIAGA 292
DB 333 NVPPGADEIMINKPYRSATGTSMTAPMVAGAVALLLE---KNSRLTNVE--IKNLIKTTA 387
QY 293 T---DVGLGYGPGQGWGRVTLDKSL 315
DB 388 TKINEAGL-----WTQGSQMINIEAL 409

RESULT 15
Q8GCT4
AC Q8GCT4; PRELIMINARY; PRT; 1237 AA.
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Subtilisin-like secreted protease.
OS Streptomyces atroolivaceus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_taxid=66869;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22336326; PubMed=12446651;
RA Cheng Y.Q., Tang G.L., Shen B.;
RT "Identification and Localization of the Gene Cluster Encoding
RT Biosynthesis of the Antitumor Macrolactam Leinamycin in Streptomyces
RT atroolivaceus S-140."
RL J. Bacteriol. 184:7013-7024 (2002).
DR EMBL; AF484556; AAN85481.1; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF02225; PA; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.

```

```

DR PROSITE; PS00137; SUBTILASE HIS; 1.
DR PROSITE; PS00138; SUBTILASE SER; 1.
DR Protease.
KW SEQUENCE 1237 AA; 129488 MW; 266C2B2760223D40 CRC64;

Query Match      15.3%; Score 346.5; DB 2; Length 1237;
Best Local Similarity 30.4%; Pred. No. 7.4e-12;
Matches 140; Conservative 55; Mismatches 154; Indels 111; Gaps 23;

QY 6 GIVKADVAQN-----YGLYGQGVVAVADTGLDTRNDSSMHEAFRGKI--TAL 53
DB 213 GRVKADLADSTAQIGAKVWAEGDTGQDVKVAWLDGADT-----EHPDLVGQVSDSAS 266
QY 54 YALGRTNNANDPENGEGTHVAGSVLG-----NALNKGMAPQANLVFQSIMDSSGGLGLPS 108
DB 267 FVPEGDDIA-DYNGEGTHVASTIVGTGSASDGKRGVAGSARLSVGKVLNSEG--SGQES 323
QY 109 NLNTLFSQAWNA---GARIHTNSWGAPVNGAYTAN---SRQVDEYVRNNDMTVLPAAGNE 162
DB 324 WI--TAGMEWAARDQKARIISMSLG---GGGDKNDPMQAVDELSDHTGALFVIAAGNG 377
QY 163 GENSGETISAPGTAKNAITVGATENYRPSFGSIADPNPHIAQFSSRG-ATRDGRIKPDVTA 221
DB 378 GPHS--ISSPGAADGALTVGA-----VDSDTLADFSSQGPDRDGDGLKPEITA 424
QY 222 PGTFLSARSSLAPDSSFWANNYSKYANGGTSMTPIVAGNVAQLREHFIKNRGITPKP 281
DB 425 PGVDIVAARSHYKRGSGY-----YTTMSGTSMATPHVAGVAALLAAAEHPDWTGTOLKE 477
QY 282 SLIKKALITAGATDVGLGYPSGQGWGRVTLDKSLNAVYVNEATALTTCOKATYSFQTQAG 341
DB 478 ALVSSAKATPA-----YTPYQAGAGRLDAPAAVHTTVPATTAYS-----GHEH 521
QY 342 KPLKISLVMTDAPGST---TASYTLVND-----LDLVI--TAPNGOKYVGNDFSPYDNNW 392
DB 522 -----WPPKPGETDVRTVTYTNVGDAPVSLNLAVNGTVPAGL-----FSLSEHDH-- 565
QY 393 DGRNNVENVFINAPSGTYTTEV-QAYNVPSGPORFSLAI 431
DB 566 -----VIVPAHGATVTTLTAALOKLAGDQSVSAVI 595

```

Search completed: March 31, 2004, 16:08:57
Job time : 33.475 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:03 ; Search time 47.1304 Seconds
(without alignments)
2595.843 Million cell updates/sec

Title: US-09-985-689A-5
Perfect score: 2261
Sequence: 1 NDVARGIVKADVAQNNGLY.....EVQAYNVPSQGRFLAIVH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2261	100.0	433	5	AAM50084 Bacillus
2	2251	99.6	433	5	AAM50082 Bacillus
3	2245	99.3	433	2	AAR26274 Alkali-pr
4	2245	99.3	433	2	AAM61495 Modified
5	2245	99.3	433	2	AAM95698 Bacillus
6	2245	99.3	433	3	AAY69207 Amino aci
7	2245	99.3	433	3	AAY44619 Bacillus
8	2240	99.1	433	5	AAM50083 Bacillus
9	2236	98.9	433	5	AAM89548 Bacillus
10	2044	90.4	433	5	AAM50086 Bacillus
11	2044	90.4	433	5	AAM89547 Bacillus
12	2024	89.5	434	5	AAM50085 Bacillus
13	2010	88.9	434	5	AAM50081 Bacillus
14	2002	88.6	439	2	AAY17089 Bacillus
15	1999	88.4	434	5	AAY17091 Bacillus
16	1998	88.4	434	5	AAM50080 Bacillus
17	1998	88.4	640	2	AAY17090 Bacillus
18	1952	86.4	639	2	AAY17087 An alkali
19	1952	86.4	640	2	AAY17088 An alkali
20	1841	81.4	434	5	AAM50090 Bacillus
21	1514	67.0	345	2	AAM62230 Subtilase
22	1514	67.0	345	2	AAY21654 Subtilase
23	443	19.6	659	2	AAM24121 Thermococ
24	443	19.6	659	2	AAM94840 WO9856926
25	422	18.7	412	2	AAM94836 Hyperther

26	422.5	18.7	522	2	AAW24122	Aaw24122 Pyrococcus
27	422.5	18.7	522	2	AAW94838	Aaw94838 Hyperther
28	422.5	18.7	654	2	AAW24129	Aaw24129 Pyrococcus
29	422.5	18.7	654	2	AAW94841	Aaw94841 Hyperther
30	401.5	17.8	659	2	AAW24123	Aaw24123 Protease.
31	373.5	16.5	545	4	ABSO9483	ABSO9483 T. yonsei
32	347.5	15.4	1079	4	AAW1180	AAW1180 Transglut
33	347.5	15.4	1079	6	ABU07391	ABU07391 Foreign p
34	343	15.2	520	2	AAW13666	AAW13666 Fragment
35	343	15.2	734	2	AAW13667	AAW13667 Streptomy
36	343	15.2	823	2	AAW13668	AAW13668 DhpA-mel
37	341.5	15.1	1237	6	ABU11343	ABU11343 Protein e
38	342	13.8	806	2	AAR27481	AAR27481 RP-III re
39	307	13.6	903	2	AAR87007	AAR87007 Hyperther
40	307	13.6	1398	2	AAR87008	AAR87008 Protease.
41	307	13.6	1398	2	AAW24124	AAW24124 Pyrococcus
42	307	13.6	1398	2	AAW94839	Aaw94839 WO9856926
43	298.5	13.2	580	7	ADD24927	Add24927 Xanthomon
44	295	13.0	519	6	ABP76735	ABP76735 Streptomy
45	295	13.0	19938	6	ABP76678	ABP76678 Streptomy

ALIGNMENTS

RESULT 1
AAM50084
ID AAM50084 standard; protein; 433 AA.

XX AC AAM50084;
XX AC
DT 12-AUG-2002 (first entry)
XX DE Bacillus sp SD-521 (FERM BP-11162) alkaline protease protein fragment.
XX KW Alkaline protease; detergent; laundry; bleaching; dishwasher.
XX OS Bacillus sp.
XX PN EPI2092333-A2.
XX PD 29-MAY-2002.
XX PF 22-NOV-2001; 2001EP-00127851.
XX PR 22-NOV-2000; 2000JP-00355166.
XX PR 12-APR-2001; 2001JP-00114048.
XX FA (KAOS) KAO CORP.
XX FI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
XX FI Okuda M, Saeki K;
XX DR WPI; 2002-437518/47.
XX PT New modified alkaline proteases useful in detergent compositions.
XX PS Claim 5; Page 16-18; 25pp; English.
XX CC This invention describes novel Bacillus sp. alkaline proteases useful in
XX CC detergent compositions, especially in laundry, bleaching or automatic
XX CC dishwasher detergents. The novel proteases have an increased detergency &
XX CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX CC sequence represents a fragment of the alkaline protease SD-521 from
XX CC Bacillus sp strain SD-521 (FERM BP-11162) described in the method of the
XX CC invention
SQ Sequence 433 AA;

Query Match 100.0%; Score 2261; DB 5; Length 433;
Best Local Similarity 100.0%; Pred. No. 3.8e-168;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 NDVARGIVKADVQAQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
DB 1 NDVARGIVKADVQAQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
DB 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGPNSTGISAPGTAKNAIT 180
DB 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGPNSTGISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSLADPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSSLAPDSSEW 240
DB 181 VGATENYRPSFGSLADPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSSLAPDSSEW 240
QY 181 VGATENYRPSFGSLADPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSSLAPDSSEW 240
DB 181 VGATENYRPSFGSLADPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSSLAPDSSEW 240
QY 241 ANYNSKIAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 300
DB 241 ANYNSKIAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 300
QY 301 SGDQGWGRVTLDKSLNVAVYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
DB 301 SGDQGWGRVTLDKSLNVAVYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
QY 361 YTLVNDLDELIVITAPNGQKYGNDFSYPYDNNWGRNNVFNAPQSGTYTIEVQAYNV 420
DB 361 YTLVNDLDELIVITAPNGQKYGNDFSYPYDNNWGRNNVFNAPQSGTYTIEVQAYNV 420
QY 421 PSGQRFSLAIVH 433
DB 421 PSGQRFSLAIVH 433

```

RESULT 2

AAW50082 standard; protein; 433 AA.

XX AC AAW50082;

XX 12-AUG-2002 (first entry)

XX Bacillus sp D6-(FERM P1592) alkaline protease protein fragment.

XX Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX Bacillus sp.

XX EP1209233-A2.

XX 29-MAY-2002.

XX 22-NOV-2001; 2001EP-00127851.

XX 22-NOV-2000; 2000JP-00355166.

XX 12-APR-2001; 2001JP-00114048.

XX (XAOS) KAO CORP.

XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;

XX Okuda M, Saeki K;

XX WPI; 2002-437518/47.

XX New modified alkaline proteases useful in detergent compositions.

XX Claim 5; Page 13-15; 25pp; English.

XX This invention describes novel Bacillus sp. alkaline proteases useful in
 CC detergent compositions, especially in laundry, bleaching or automatic
 CC dishwasher detergents. The novel proteases have an increased detergent *
 CC (34 - 38) compared to prior art alkaline proteases (31 and 23). This
 CC sequence represents a fragment of the alkaline protease E-1 from Bacillus
 CC sp strain D6-(FERM-P1592) described in the method of the invention

XX SQ Sequence 433 AA;

Query Match 99.6%; Score 2251; DB 5; Length 433;

Best Local Similarity 99.3%; Pred. No. 2.3e-167;

Matches 430; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 NDVARGIVKADVQAQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
DB 1 NDVARGIVKADVQAQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
DB 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGPNSTGISAPGTAKNAIT 180
DB 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFPAAGNEGPNSTGISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSLADPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSSLAPDSSEW 240
DB 181 VGATENYRPSFGSLADPNHIAQFSSRGATRGRIKPDVTAPGTFFILSARSSSLAPDSSEW 240
QY 241 ANYNSKIAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 300
DB 241 ANYNSKIAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGYP 300
QY 301 SGDQGWGRVTLDKSLNVAVYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
DB 301 SGDQGWGRVTLDKSLNVAVYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
QY 361 YTLVNDLDELIVITAPNGQKYGNDFSYPYDNNWGRNNVFNAPQSGTYTIEVQAYNV 420
DB 361 YTLVNDLDELIVITAPNGQKYGNDFSYPYDNNWGRNNVFNAPQSGTYTIEVQAYNV 420
QY 421 PSGQRFSLAIVH 433
DB 421 PSGQRFSLAIVH 433

```

RESULT 3

AAW50082 standard; protein; 433 AA.

XX AC AAW50082;

XX 05-FEB-1993 (first entry)

XX Alkali-protease Ya enzyme.

XX Alkali resistance; surface active agent resistance; detergent improver.

XX Bacillus sp. Y.

XX JP04197182-A.

XX 16-JUL-1992.

XX 28-NOV-1990; 90JP-00327110.

XX 28-NOV-1990; 90JP-00327110.

XX (LIOY) LION CORP.

XX WPI; 1992-288440/35.

XX N-PSDB; AAQ27516.

XX DNA coding alkali-protease Ya enzyme - has good alkali and surfactant
 PT resistance and improves detergent.

XX Claim 2; Page 1; 17pp; Japanese.

XX The sequence is that of alkali-protease Ya enzyme which can be used in

CC the recombinant production of Ya enzyme. Ya enzyme is excellent in alkali
 CC resistance and surface active agent resistance and improves detergent
 XX
 SQ Sequence 433 AA;

Query Match 99.3%; Score 2245; DB 2; Length 433;
 Best Local Similarity 99.1%; Pred. No. 6.8e-167;
 Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSTMDSSGGLGGLPSNLTLFSQAWNA 120
 DB 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSTMDSSGGLGGLPSNLTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSLADNPNHIAQFSRGATRDGRIPKPDVTAPGTFTLSARSLAPDSFSW 240
 DB 181 VGATENYRPSFGSLADNPNHIAQFSRGATRDGRIPKPDVTAPGTFTLSARSLAPDSFSW 240
 QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
 DB 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
 QY 301 SGDOGWGRVTLDKSLNVAYVNEATATGOKATYSFOAQAGKPLKISLVWTDAGSGTTAS 360
 DB 301 SGDOGWGRVTLDKSLNVAYVNEATATGOKATYSFOAQAGKPLKISLVWTDAGSGTTAS 360
 QY 361 YTLVNDLDLVIAPNGQKYVGNDFSYPDYNNWDRNENVENFINAPQSGTIIIEVQAYNV 420
 DB 361 YTLVNDLDLVIAPNGQKYVGNDFSYPDYNNWDRNENVENFINAPQSGTIIIEVQAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433

RESULT 4
 AA61495
 ID AA61495 standard; protein; 433 AA.
 AC AA61495;
 XX
 XX
 DT 06-NOV-1998 (first entry)
 DE Modified Bacillus lion Y protease.
 KW Bacillus lion Y protease; polyethylene glycol; PEG; soap;
 KW methoxypolyethyleneglycol; mPEG; skin; hair care product; cosmetic;
 KW lipstick; hair gel; sun oil; shampoo; hair dye; insect repellent.
 OS Bacillus sp.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1..433
 FT /note= "The enzyme is modified by methoxypolyethyleneglycol
 FT molecules covalently attached to the N-terminal amino
 FT group and to fourteen unspecified amino groups of lysine
 FT residues present on the surface of the enzyme"
 XX
 FN W09830682-A1.
 XX
 XX
 PD 16-JUL-1998.
 XX
 XX
 PF 12-JAN-1998; 98WO-DK000015.
 XX
 XX
 PR 10-JAN-1997; 97DK-00000038.
 PR 25-JUN-1997; 97DK-00000754.

XX
 PA (NOVO) NOVO-NORDISK AS.

XX
 PI Olsen AA, Pronto A;

XX
 DR WPI; 1998-399132/34.

XX
 PT New enzyme modified by attachment of many polymeric molecules - useful in
 PT skin and hair care products, has reduced tendency to cause sensitisation
 PT and increased stability.

XX
 PS Claim 16; Page 44-45; 56pp; English.

XX
 CC The present sequence represents the Bacillus lion Y protease. The
 CC invention claims for enzymes covalently modified on their surface by the
 CC attachment of a large number of small polymeric molecules, e.g.
 CC polyethylene glycols (PEG). The polymeric molecules are coupled to the N-
 CC terminal amino group and the amino groups of lysine residues found on the
 CC surface of the enzyme. In the example given, the lion Y protease was
 CC modified using methoxypolyethyleneglycol (mPEG) as the polymeric molecule.
 CC The N-terminal amino group and the amino groups of the fourteen lysine
 CC residues present on the surface of the lion Y protease were modified.
 CC Modification of the enzymes increases the stability and/or reduces the
 CC sensitising potential (allergenicity) of the enzyme, without
 CC significantly reducing enzymatic activity. Also, using a large number of
 CC relatively small polymeric molecules, rather than a few very large ones,
 CC provides a more even effect with reduced activity loss. The modified
 CC enzymes are claimed to be useful as components of a wide range of skin
 CC and hair care products, e.g. soaps, cosmetics, creams, gels, lipsticks,
 CC hair gels, sun oils, shampoos, hair dyes, insect repellants, etc

XX
 SQ Sequence 433 AA;

Query Match 99.3%; Score 2245; DB 2; Length 433;
 Best Local Similarity 99.1%; Pred. No. 6.8e-167;
 Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSTMDSSGGLGGLPSNLTLFSQAWNA 120
 DB 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSTMDSSGGLGGLPSNLTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAPVNGAYTANSROVDYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSLADNPNHIAQFSRGATRDGRIPKPDVTAPGTFTLSARSLAPDSFSW 240
 DB 181 VGATENYRPSFGSLADNPNHIAQFSRGATRDGRIPKPDVTAPGTFTLSARSLAPDSFSW 240
 QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
 DB 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
 QY 301 SGDOGWGRVTLDKSLNVAYVNEATATGOKATYSFOAQAGKPLKISLVWTDAGSGTTAS 360
 DB 301 SGDOGWGRVTLDKSLNVAYVNEATATGOKATYSFOAQAGKPLKISLVWTDAGSGTTAS 360
 QY 361 YTLVNDLDLVIAPNGQKYVGNDFSYPDYNNWDRNENVENFINAPQSGTIIIEVQAYNV 420
 DB 361 YTLVNDLDLVIAPNGQKYVGNDFSYPDYNNWDRNENVENFINAPQSGTIIIEVQAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433

RESULT 5
 AA61495
 ID AA61495 standard; protein; 433 AA.

XX AAW95698;
 AC
 XX
 DT 16-JUN-1999 (first entry)
 XX
 DE Bacillus sp. Lion Y protease.
 XX
 KW PD498; subtilisin; Lion Y; protease; skin-care; feed; additive; soap;
 KW cosmetic; hair dye; sunscreens; acne; antiperspirants; insect repellent;
 KW deodorant; detergent; food; breadmaking; textile-treating; oral; dermal;
 KW pharmaceutical; agrochemical.
 XX
 OS Bacillus sp.
 XX
 PN WO9900489-A1.
 XX
 XX 07-JAN-1999.
 XX
 XX 22-JUN-1998; 98WO-DK000270.
 XX
 XX 25-JUN-1997; 97DK-00000753.
 PR 07-JUL-1997; 97US-0051830P.
 XX
 XX (NOVO) NOVO-NORDISK AS.
 PA
 XX Olsen AA, Fatum TM, Deussen H, Roggen EL;
 PI WPI; 1999-095735/08.
 XX
 XX New modified polypeptide with attached low molecular weight polymer - has
 PT reduced respiratory allergenicity, useful in skin care products,
 PT detergents, as food additives or textile-treating compositions.
 PT
 XX Claim 10; Page 48-49; 60pp; English.
 PS
 XX The sequence is that of Lion Y protease. This can be used as an active
 CC ingredient: (i) in personal care products (especially skin-care products
 CC such as soaps, cosmetics, hair dyes, sunscreens, anti-acne products,
 CC antiperspirants, insect repellants or deodorants); (ii) in detergents (as
 CC laundry, dishwashing or hard-surface cleaners); (iii) food or feed
 CC additives (e.g. for breadmaking); (iv) in textile-treating compositions,
 CC or (v) in oral or dermal pharmaceuticals and agrochemicals
 XX
 SQ Sequence 433 AA;
 Query Match 99.3%; Score 2245; DB 2; Length 433;
 Best Local Similarity 99.1%; Pred. No. 6.8e-167;
 Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAGQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAGQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 DB 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 DB 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
 QY 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIPKPDVTAPCTFILSARSSLAPDSSFW 240
 DB 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIPKPDVTAPCTFILSARSSLAPDSSFW 240
 QY 241 ANYNSKYAYMGGTSMATFIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVLGYVP 300
 DB 241 ANYNSKYAYMGGTSMATFIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVLGYVP 300
 QY 301 SGDCGWRVTLDSLNVAAYNEATATGOKATYSFQAQAGKPKISLWWTDPAGSTTAS 360
 DB 301 SGDCGWRVTLDSLNVAAYNEATATGOKATYSFQAQAGKPKISLWWTDPAGSTTAS 360

QY 361 YTLVNDLDELVTAPNGQKYVGNDFSPYDNNNDGNNVNFINAPQSGTYTIEVOAYNV 420
 DB 361 YTLVNDLDELVTAPNGQKYVGNDFSPYDNNNDGNNVNFINAPQSGTYTIEVOAYNV 420
 QY 421 PSGQRFSLAIVH 433
 DB 421 PSGQRFSLAIVH 433
 RESULT 6
 AAY69207
 ID AAY69207 standard; protein; 433 AA.
 XX
 AC AAY69207;
 XX
 DT 30-MAY-2000 (first entry)
 XX
 DE Amino acid sequence of protease Lion Y.
 XX
 KW Protease Lion Y; polypeptide-polymer conjugate; washing performance;
 KW respiratory allergenicity; allergic reaction; detergent formulation;
 KW laundry; dishwashing; hard surface cleaner; agricultural chemical;
 KW skin care; cosmetic; oral pharmaceutical; dental pharmaceutical;
 KW textile processing.
 XX
 OS Bacillus sp.
 XX
 PN WO200004138-A1.
 XX
 XX 27-JAN-2000.
 PD
 XX 16-JUL-1999; 99WO-DK000406.
 PF
 XX 17-JUL-1998; 98DK-00000951.
 PR
 XX (NOVO) NOVO-NORDISK AS.
 PA
 XX Bauditz P, Fatum TM, Olsen AA, Deussen H, Petersen DA;
 PI WPI; 2000-195024/17.
 XX
 XX New polypeptide-polymer conjugate, particularly enzyme conjugate, useful
 PT in detergent formulations.
 PT
 XX Disclosure; Page 74-76; 79pp; English.
 PS
 XX The present sequence represents a protease Lion Y protein. The protein
 CC may be used to produce the conjugates of the invention. The specification
 CC describes polypeptide-polymer conjugates which have improved washing
 CC performance and reduced respiratory allergenicity compared with the
 CC unconjugated polypeptide. The polymer provides, in water, a conformation
 CC that shields the molecular surface effectively, preventing association of
 CC antibodies that can induce an allergic reaction. The conjugates are used
 CC in industrial compositions, particularly detergent formulations (laundry,
 CC dishwashing or hard surface cleaners), but also in agricultural
 CC chemicals, skin care products (cosmetics and toiletries), oral and dental
 CC pharmaceuticals, or textile processing and treatment compositions
 XX
 SQ Sequence 433 AA;
 Query Match 99.3%; Score 2245; DB 3; Length 433;
 Best Local Similarity 99.1%; Pred. No. 6.8e-167;
 Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NDVARGIVKADVAGQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 DB 1 NDVARGIVKADVAGQNNYGLYGQGVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 DB 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLFSQAWNA 120
 QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180

```

Db 121 GARIHNSWGAPVNGAYTANSRQVDEYVRNDMTVLFAGNEGPNSTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGLADPNPHIAQFSRGATRDGRIRKPDVTPGTFTLSARSLAPDSSFW 240
Db 181 VGATENYRPSFGLADPNPHIAQFSRGATRDGRIRKPDVTPGTFTLSARSLAPDSSFW 240
QY 241 ANYNSKYAYMGTSKATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVLGY 300
Db 241 ANYNSKYAYMGTSKATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVLGY 300
QY 301 SGQGWGRVTLDKSLNVAYNEATATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
Db 301 SGQGWGRVTLDKSLNVAYNEATATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
QY 361 YTLVNDLDELVTAPNGQKYVGNDFSYPDNNWDGRNNVENVFINAPQSGTYTIEVOAYNV 420
Db 361 YTLVNDLDELVTAPNGQKYVGNDFSYPDNNWDGRNNVENVFINAPQSGTYTIEVOAYNV 420
QY 421 PSGPQRFSLAIVH 433
Db 421 PSGPQRFSLAIVH 433

```

RESULT 7

AA44619
ID AAY44619 standard; protein; 433 AA.

XX AAY44619;
DT 07-APR-2000 (first entry)
XX Bacillus Lion Y enzyme.

XX Lion Y enzyme; protease; allergic response; industrial composition;
KW co-polymer; ethylene oxide; EO; propylene oxide; PO; conjugate;
KW allergenicity; detergent; cosmetic; toiletries; textile treatment;
KW agrochemical; pharmaceutical; food; feed additive.

XX Bacillus sp.

XX WO9967370-A1.

XX 29-DEC-1999.

XX 23-JUN-1999; 99WO-DK000359.

XX 23-JUN-1998; 98DX-00000809.

XX (NOVO) NOVO-NORDISK AS.

XX Deussen H, Olsen AA, Fatum TM, Roggen EL;

XX WPI; 2000-136981/12.

XX New conjugate of polypeptide, especially an enzyme, with copolymer of
PT ethylene oxide and propylene oxide, used in e.g. cleaning compositions,
PT has reduced allergenicity.

PS Claim 10; Page 56-57; 62pp; English.

XX The present sequence is a Bacillus Lion Y enzyme, which is a protease
CC capable of inducing an allergic response upon inhalation. The enzyme can
CC be covalently coupled to a co-polymer comprising ethylene oxide (EO) and
CC propylene oxide (PO) to reduce its allergenicity. This enzyme-polymer
CC conjugate can be used in industrial compositions such as detergents,
CC cosmetics, toiletries, textile treatment compositions, agrochemicals,
CC oral and dermal pharmaceuticals and food and feed additives

XX Sequence 433 AA;

Query Match 99.3%; Score 2245; DB 3; Length 433;

Best Local Similarity 99.1%; Pred. No. 6.8e-167;

```

Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNNGYLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQNNGYLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
QY 61 NNDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIWDSGGGLGGLPSNLTLFSQAWNA 120
Db 61 NNDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIWDSGGGLGGLPSNLTLFSQAWNA 120
QY 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNDMTVLFAGNEGPNSTISAPGTAKNAIT 180
Db 121 GARIHTNSWGAPVNGAYTANSRQVDEYVRNDMTVLFAGNEGPNSTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGLADPNPHIAQFSRGATRDGRIRKPDVTPGTFTLSARSLAPDSSFW 240
Db 181 VGATENYRPSFGLADPNPHIAQFSRGATRDGRIRKPDVTPGTFTLSARSLAPDSSFW 240
QY 241 ANYNSKYAYMGTSKATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVLGY 300
Db 241 ANYNSKYAYMGTSKATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVLGY 300
QY 301 SGQGWGRVTLDKSLNVAYNEATATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
Db 301 SGQGWGRVTLDKSLNVAYNEATATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
QY 361 YTLVNDLDELVTAPNGQKYVGNDFSYPDNNWDGRNNVENVFINAPQSGTYTIEVOAYNV 420
Db 361 YTLVNDLDELVTAPNGQKYVGNDFSYPDNNWDGRNNVENVFINAPQSGTYTIEVOAYNV 420
QY 421 PSGPQRFSLAIVH 433
Db 421 PSGPQRFSLAIVH 433

```

RESULT 8

AA50083
ID AAM50083 standard; protein; 433 AA.

XX AAM50083;

DT 12-AUG-2002 (first entry)

XX Bacillus sp Y-(FERM BP-1029) alkaline protease protein fragment.

XX Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX Bacillus sp.

XX EPI209233-A2.

XX 29-MAY-2002.

XX 22-NOV-2001; 2001EP-00127851.

XX 22-NOV-2000; 2000JP-00355166.

XX 12-APR-2001; 2001JP-00114048.

XX (KAOS) KAO CORP.

XX Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
PI Okuda M, Saeki K;

XX WPI; 2002-437518/47.

XX New modified alkaline proteases useful in detergent compositions.

PS Claim 5; Page 15-16; 25pp; English.

XX This invention describes novel Bacillus sp. alkaline proteases useful in
CC detergent compositions, especially in laundry, bleaching or automatic
CC dishwasher detergents. The novel proteases have an increased detergency &
CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This

CC sequence represents a fragment of the alkaline protease Ya from *Bacillus*
 CC sp strain Y-(FERM BP-1029) described in the method of the invention

XX Sequence 433 AA;

Query Match 99.1%; Score 2240; DB 5; Length 433;

Best Local Similarity 98.8%; Pred. No. 1.7e-166; Indels 0; Gaps 0;

Matches 428; Conservative 4; Mismatches 1;

QY 1 NDVARGIVKADVAQNNGYGLGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60

DB 1 NDVARGIVKADVAQNNGYGLGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60

QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLSQAWNA 120

DB 61 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLSQAWNA 120

QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAIT 180

DB 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAIT 180

QY 181 VGATENYRPSFGLADPNPHIAQFSSRGATRGRIKPDVTAFGTILSARSSLAPDSSF 240

DB 181 VGATENYRPSFGLADPNPHIAQFSSRGATRGRIKPDVTAFGTILSARSSLAPDSSF 240

QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300

DB 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300

QY 301 SGQGWGRVTLDKSLNVAVNEATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 360

DB 301 NGDQGWGRVTLNKLNSVAYNEATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 360

QY 361 YTLVNDLDELVTAPNGQKYGNDFSYPYDNNWDRNNVNFVFINAPQSGTYTIEVQAYNV 420

DB 361 YTLVNDLDELVTAPNGQKYGNDFSYPYDNNWDRNNVNFVFINAPQSGTYTIEVQAYNV 420

QY 421 PSQPQRFSLAIVH 433

DB 421 PSQPQRFSLAIVH 433

RESULT 9

AAW89548

ID AAW89548 standard; protein; 636 AA.

AC AAW89548;

XX 12-APR-1999 (first entry)

DE *Bacillus* sp. alkaline protease Y.

KW Alkaline protease Y; detergent; surfactant; leather processing;

XX debittering; flavour.

OS *Bacillus* sp.

XX WO9856927-A2.

XX 17-DEC-1998.

XX 09-JUN-1998; 98WO-US012005.

XX 12-JUN-1997; 97US-00873479.

XX (NOVO) NOVO NORDISK BIOTECH INC.

XX Sloma A, Christianson L;

XX WPI; 1999-080908/07.

XX Novel protease from *Bacillus subtilis* LC20 - useful in laundry and

PT dishwashing detergents and for leather processing.

XX

PS Claim 3; Page 55-56; 77pp; English.

XX

CC This is the amino acid sequence of a *Bacillus* sp. alkaline protease Y that is said to have good alkali and surfactant resistance and improved detergency. It shows 77% identity to a newly isolated protease (see AAW89547) of *Bacillus* sp. JP170 (NCIB 12513). The invention provides vectors, recombinant host cells and methods for the recombinant production of such proteases. The protease are used in laundry and dishwashing detergents, for institutional and industrial cleaning, and for leather processing, as well as for debittering and enhancing the degree of hydrolysis of protein hydrolysates, for flavour development and through hydrolysis of proteins, degradation of undesired peptides and enzymatic synthesis of peptides. They have enhanced stability towards oxidation under alkaline conditions, e.g. towards bleaching agents of the proxy type. The invention also provides mutant cells in which the protease activity is diminished. Such cells can be used for the production of heterologous recombinant proteins

XX Sequence 636 AA;

Query Match 98.9%; Score 2236; DB 2; Length 636;

Best Local Similarity 98.8%; Pred. No. 5.9e-166;

Matches 428; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYGLGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60

DB 204 NDVARGIVKADVAQNNGYGLGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 263

QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLSQAWNA 120

DB 264 NASDPNGHGHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLTLSQAWNA 323

QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAIT 180

DB 324 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAIT 383

QY 181 VGATENYRPSFGLADPNPHIAQFSSRGATRGRIKPDVTAFGTILSARSSLAPDSSF 240

DB 384 VGATENYRPSFGLADPNPHIAQFSSRGATRGRIKPDVTAFGTILSARSSLAPDSSF 443

QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300

DB 444 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 503

QY 301 SGQGWGRVTLDKSLNVAVNEATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 360

DB 504 SGQGWGRVTLDKSLNVAVNEATATGOKATYSFOAQAGKPLKISLWTDAPGSTTAS 563

QY 361 YTLVNDLDELVTAPNGQKYGNDFSYPYDNNWDRNNVNFVFINAPQSGTYTIEVQAYNV 420

DB 564 YTLVNDLDELVTAPNGQKYGNDFSYPYDNNWDRNNVNFVFINAPQSGTYTIEVQAYNV 623

QY 421 PSQPQRFSLAIVH 433

DB 624 PSQPQRFSLAIVH 636

RESULT 10

AAW50086

ID AAW50086 standard; protein; 433 AA.

AC AAW50086;

XX 12-AUG-2002 (first entry)

DE *Bacillus* sp. alkaline protease protein A-2 fragment.

KW Alkaline protease; detergent; laundry; bleaching; dishwasher.

XX *Bacillus* sp.

XX EP1209233-A2.

```

XX PD 29-MAY-2002.
XX PF
XX PF 22-NOV-2001; 2001EP-00127851.
XX PR
XX PR 22-NOV-2000; 2000JP-00355166.
XX PR 12-APR-2001; 2001JP-00114048.
XX PA
XX PA (KAOS ) KAO CORP.
XX PI
XX PI Hatada Y, Ogawa A, Kageyama Y, Sato T, Araki H, Sumitomo N;
XX PI Okuda M, Saeki K;
XX XX
XX DR WPI; 2002-437518/47.
XX XX
XX PT New modified alkaline proteases useful in detergent compositions.
XX PS
XX PS Claim 5; Page 20-21; 25pp; English.
XX CC
XX CC This invention describes novel Bacillus sp. alkaline proteases useful in
XX CC detergent compositions, especially in laundry, bleaching or automatic
XX CC dishwasher detergents. The novel proteases have an increased detergency &
XX CC (34 - 38%) compared to prior art alkaline proteases (31 and 23%). This
XX CC sequence represents a fragment of the alkaline protease A-2 from Bacillus
XX CC sp NCIB12513 described in the method of the invention
XX XX
XX SQ Sequence 433 AA;
XX
Query Match 90.4%; Score 2044; DB 5; Length 433;
Best Local Similarity 89.1%; Pred. No. 3.5e-151;
Matches 386; Conservative 24; Mismatches 23; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAAQNNYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
DB 1 NDVARGIVKADVAAQNNYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAQANLVFQSIIMDSGGGLGGLPANLTLFQAWNA 120
DB 61 NANDPNHGHTHVAGSVLGNALNKGMAQANLVFQSIIMDSGGGLGGLPANLTLFQAWNA 120
QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
DB 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
QY 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRIKPDVTAPGTITLSARSSSLAPDSFW 240
DB 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRIKPDVTAPGTITLSARSSSLAPDSFW 240
QY 241 ANYNSKYAYNGGTSMATPIVAGNVAGLREHFVKRGITKPSLIKAALIAGATDYLGY 300
DB 241 ANYNSKYAYNGGTSMATPIVAGNVAGLREHFVKRGITKPSLIKAALIAGATDYLGY 300
QY 301 SGQDQWGRVTLDKSLNVAVYNEATATGOKATYSFQAQAGKPLKISLVWTDAPGSTTAS 360
DB 301 NGNQWGRVTLDKSLNVAVFVNETSPSTSQATYSFTAQAGKPLKISLVWSDAPGSTTAS 360
QY 361 YTLVNDLVLITAPNGOKYVNDPSPYDNNMDGRNNVENFINAPQSGTYTIEVOAYNV 420
DB 361 LTLVNDLVLITAPNGTKYVNDFTAPYDNNMDGRNNVENFINAPQSGTYTIEVOAYNV 420
QY 421 PSQPQRFSLAIVH 433
DB 421 PVSPQTFSLAIVH 433
XX
RESULT 11
ID AAW89547
XX AC AAW89547 standard; protein; 641 AA.
XX AC AAW89547;
XX XX
XX DT 12-APR-1999 (first entry)
XX XX

```

```

DE DE Bacillus JPI70 protease.
XX XX
XX KW Protease; detergent; surfactant; leather processing; debittering;
XX KW flavour.
XX OS
XX OS Bacillus sp.
XX PH
XX PH Key Location/Qualifiers
XX FT Peptide 1..33
XX FT /note= "signal peptide"
XX FT Region 34..208
XX FT /note= "prepro region"
XX FT Protein 209..641
XX FT /note= "mature protein"
XX XX
XX XX WO9856927-A2.
XX PN
XX PN 17-DEC-1998.
XX PD
XX PD 09-JUN-1998; 98WO-US012005.
XX PF
XX PF 12-JUN-1997; 97US-00873479.
XX PR
XX PR (NOVO ) NOVO NORDISK BIOTECH INC.
XX PA
XX PA Sloma A, Christianson L;
XX PI
XX PI WPI; 1999-080908/07.
XX DR
XX DR N-PSDB; AAV82382.
XX FT
XX FT Novel protease from Bacillus subtilis LC20 - useful in laundry and
XX FT dishwashing detergents and for leather processing.
XX PS
XX PS Claim 7; Page 53-54; 77pp; English.
XX CC
XX CC This is the amino acid sequence of a novel protease of Bacillus sp. JPI70
XX CC (NCIB 12513), as deduced from the nucleotide sequence of an isolated gene
XX CC (see AAV82382). The entire protein, including the signal peptide and
XX CC prepro region, has 77% identity to alkaline protease Y (see AAW89548)
XX CC from Bacillus. The invention provides vectors, recombinant host cells and
XX CC methods for the recombinant production of the protease. The protease is
XX CC used in laundry and dishwashing detergents, for institutional and
XX CC industrial cleaning, and for leather processing, as well as for
XX CC debittering and enhancing the degree of hydrolysis of protein
XX CC hydrolysates, for flavour development through hydrolysis of proteins,
XX CC degradation of undesired peptides and in enzymatic synthesis of peptides.
XX CC It has enhanced stability towards oxidation under alkaline conditions,
XX CC e.g. towards bleaching agents of the peroxy type. The invention also
XX CC provides mutant cells in which the protease activity is diminished. Such
XX CC cells can be used for the production of heterologous recombinant proteins
XX SQ Sequence 641 AA;
XX
Query Match 90.4%; Score 2044; DB 2; Length 641;
Best Local Similarity 89.1%; Pred. No. 6.1e-151;
Matches 386; Conservative 24; Mismatches 23; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAAQNNYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 60
DB 209 NDVARGIVKADVAAQNNYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALYALGRTN 269
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAQANLVFQSIIMDSGGGLGGLPANLTLFQAWNA 120
DB 269 NANDPNHGHTHVAGSVLGNALNKGMAQANLVFQSIIMDSGGGLGGLPANLTLFQAWNA 328
QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
DB 329 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 388
QY 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRIKPDVTAPGTITLSARSSSLAPDSFW 240
DB 389 VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRIKPDVTAPGTITLSARSSSLAPDSFW 448

```


Db 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 QY 61 NANDPNHGHTHVAGSVLGN-ALNKGVAPOANLVFOSIMDSSGGLGGLPSNLTLFSAWN 119
 Db 61 NANDPNHGHTHVAGSVLGN-ALNKGVAPOANLVFOSIMDSSGGLGGLPSNLTLFSAWN 120
 QY 120 AGARIHTNSWGAAPVNGAYTANRQVDEYVRNNDMTLVFAAGNEGPNSTISAPGTAKNAI 179
 Db 121 AGARIHTNSWGAAPVNGAYTANRQVDEYVRNNDMTLVFAAGNEGPNSTISAPGTAKNAI 180
 QY 180 TVGATENYRPSFGSLADNPNHIAQFSSRGATRDGRIKPDVAPGTFFLSARSSSLAPDSSF 239
 Db 181 TVGATENYRPSFGSLADNPNHIAQFSSRGATRDGRIKPDVAPGTFFLSARSSSLAPDSSF 240
 QY 240 WANYNSKYAYMGGTSMATPIVAGNVAQLREHFIPKRGITPKPSLKAALIAAGATDVGLGY 299
 Db 241 WANDHSDKYAYMGGTSMATPIVAGNVAQLREHFIPKRGITPKPSLKAALIAAGATDVGLGY 300
 QY 300 PSGDQGWGRVTLDKSLNVAIYNEATATGQKATYSFOAQAGKPLKISLVWTDAPGSTTA 359
 Db 301 PNGNQGWGRVTLDKSLNVAIYNEATATGQKATYSFOAQAGKPLKISLVWTDAPGSTTA 360
 QY 360 SVTLVNDLVLITAPNGQKYVGNDFSYPDNNWGRNENNVFINAPQSGTYTIEVQAYN 419
 Db 361 SVTLVNDLVLITAPNGQKYVGNDFSYPDNNWGRNENNVFINAPQSGTYTIEVQAYN 420
 QY 420 VPSPQRFSLAIVH 433
 Db 421 VPVGPQNFSLAIVN 434

RESULT 14

AAV17089
 ID AAV17089 standard; protein; 639 AA.

XX AAV17089;
 AC
 DT 20-MAR-2003 (revised)
 DT 21-JUL-1999 (first entry)
 XX
 DE Bacillus alkaline protease.

XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
 KW washing composition; oxidising agent.
 XX
 OS Bacillus sp.

XX
 PN WO9918218-A1.

XX
 PD 15-APR-1999.

XX
 PF 07-OCT-1998; 98WO-JP004528.

XX
 PR 07-OCT-1997; 97JP-00274570.

XX
 PA (KAOS) KAO CORP.

XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
 PI Shikata S, Nomura M;

XX
 DR WPI; 1999-287736/27.

XX
 DR N-PSDB; AAX37277.

XX Alkali protease from Bacillus used in washing powders.
 PT
 XX
 XX Disclosure; Page 53-58; 71pp; Japanese.

XX The invention relates to alkaline proteases produced by strains of
 CC Bacillus. The proteases ability to digest casein is not inhibited by
 CC oleic acid and they have a high stability to oxidising agents. The
 CC alkaline protease of the invention has the following properties: (a) it
 CC is active over the pH range 4-13 and has at least 80% of its optimum
 CC activity over the range pH 6-12; (b) after 30 minutes at 40 deg. C it is

CC stable over the pH range 6-11; c) its isoelectric point is 8.9-9.1; (d)
 CC its ability to digest casein is not inhibited by oleic acid; (e) it has
 CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
 CC used as enzymes in washing compositions for use in automatic dishwashers
 CC and for washing clothes. The stability to oxidising agents allows the
 CC enzyme to be an effective component of washing compositions including
 CC bleaches. The present sequence represents an alkaline protease. (Updated
 CC on 20-MAR-2003 to correct DR field.)

XX
 SQ Sequence 639 AA;

Query Match 88.6%; Score 2002.5; DB 2; Length 639;
 Best Local Similarity 88.2%; Pred No. 1.1e-147;
 Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
 Db 206 NDVARGIVKADVAQSSYGLYGQGIIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 265
 QY 61 NANDPNHGHTHVAGSVLGN-ALNKGVAPOANLVFOSIMDSSGGLGGLPSNLTLFSAWN 119
 Db 266 NANDPNHGHTHVAGSVLGN-ALNKGVAPOANLVFOSIMDSSGGLGGLPSNLTLFSAWN 325
 QY 120 AGARIHTNSWGAAPVNGAYTANRQVDEYVRNNDMTLVFAAGNEGPNSTISAPGTAKNAI 179
 Db 326 AGARIHTNSWGAAPVNGAYTANRQVDEYVRNNDMTLVFAAGNEGPNSTISAPGTAKNAI 385
 QY 180 TVGATENYRPSFGSLADNPNHIAQFSSRGATRDGRIKPDVAPGTFFLSARSSSLAPDSSF 239
 Db 386 TVGATENYRPSFGSLADNPNHIAQFSSRGATRDGRIKPDVAPGTFFLSARSSSLAPDSSF 445
 QY 240 WANYNSKYAYMGGTSMATPIVAGNVAQLREHFIPKRGITPKPSLKAALIAAGATDVGLGY 299
 Db 446 WANDHSDKYAYMGGTSMATPIVAGNVAQLREHFIPKRGITPKPSLKAALIAAGATDVGLGY 505
 QY 300 PSGDQGWGRVTLDKSLNVAIYNEATATGQKATYSFOAQAGKPLKISLVWTDAPGSTTA 359
 Db 506 PNGNQGWGRVTLDKSLNVAIYNEATATGQKATYSFOAQAGKPLKISLVWTDAPGSTTA 565
 QY 360 SVTLVNDLVLITAPNGQKYVGNDFSYPDNNWGRNENNVFINAPQSGTYTIEVQAYN 419
 Db 566 SVTLVNDLVLITAPNGQKYVGNDFSYPDNNWGRNENNVFINAPQSGTYTIEVQAYN 625
 QY 420 VPSPQRFSLAIVH 433
 Db 626 VPVGPQNFSLAIVN 639

RESULT 15

AAV17091
 ID AAV17091 standard; protein; 640 AA.

XX
 AC AAV17091;

XX
 DT 20-MAR-2003 (revised)

DT 21-JUL-1999 (first entry)

XX
 DE Bacillus alkaline protease.

XX Alkaline protease; Bacillus; casein digestion; oleic acid; enzyme;
 KW washing composition; oxidising agent.

XX
 OS Bacillus sp.

XX
 PN WO9918218-A1.

XX
 PD 15-APR-1999.

XX
 PF 07-OCT-1998; 98WO-JP004528.

XX
 PR 07-OCT-1997; 97JP-00274570.

XX
 PA (KAOS) KAO CORP.

XX Takaiwa M, Okuda M, Saeki K, Kubota H, Hitomi J, Kageyama Y;
PI Shikata S, Nomura M;
XX WPI; 1999-287736/27.
DR N-PSDB; AAX37279.
XX
XX Alkali protease from Bacillus used in washing powders.
XX
XX Disclosure; Page 63-68; 71pp; Japanese.
XX
XX The invention relates to alkaline proteases produced by strains of
CC Bacillus. The proteases ability to digest casein is not inhibited by
CC oleic acid and they have a high stability to oxidising agents. The
CC alkaline protease of the invention has the following properties: (a) it
CC is active over the pH range 4-13 and has at least 80% of its optimum
CC activity over the pH range 6-12; (b) after 30 minutes at 40 deg. C it is
CC stable over the pH range 6-11; (c) its isoelectric point is 8.9-9.1; (d)
CC its ability to digest casein is not inhibited by oleic acid; (e) it has
CC molecular weight about 43,000 by SDS-PAGE. The alkaline proteases can be
CC used as enzymes in washing compositions for use in automatic dishwashers
CC and for washing clothes. The stability to oxidising agents allows the
CC enzyme to be an effective component of washing compositions including
CC bleaches. The present sequence represents an alkaline protease. (Updated
CC on 20-MAR-2003 to correct DR field.)
XX
SQ Sequence 640 AA;

Query Match 88.4%; Score 1999.5; DB 2; Length 640;
Best Local Similarity 87.8%; Pred. No. 1.8e-147;
Matches 381; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVQNNYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
207 NDVARGIVKADVQNNYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 266
QY 61 NANDPNGHGTHTAGSVLGN-ALNKGMAPQANLVFQSIMDSSGGLGLPSNLTFLSQAWN 119
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
267 NANDTNGHGTHTAGSVLGNSTKNGMAPQANLVFQSIMDSSGGLGLPSNLTFLSQAYS 326
QY 120 AGARIHTNSGAPVNGAYTANSQVDEVYVNRNDMTVLFAAGNEGPNSTISAPGTAKNAI 179
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
327 AGARIHTNSGAAVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPNSTISAPGTAKNAI 386
QY 180 TVGATENYRPFSGSLADPNHIAQFSSRGATRDGRKPKDVTAPGTFTLSARSSLAPDSSF 239
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
387 TVGATENLRPSFGYADNHNHVAQFSSRGPTKDGRIKPDVWAPGTPILSARSSLAPDSSF 446
QY 240 WANYNSKYANGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGY 299
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
447 WANHDSKYANGGTSMATPIVAGNVAQLREHFVKNRGITPKPSLLKAALIAGADIGLGY 506
QY 300 PSGDQGWGRVTLDKSLNVAVYNEATATGQATYSFQAQAGKPKISLVMTDAPGSTTA 359
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
507 PNGQGWGRVTLDKSLNVAVYNESSLSLTSQATYSFTATAGKELKISLVWSDAPASTTA 566
QY 360 SYTLVNDLVLITAPNGOKYVGNDFSPYDNNWGRNNVNFINAPQSGTYTIEVQAYN 419
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
567 SVTLVNDLVLITAPNGTYVGNDFTPSYDNNWGRNNVNFINAPQSGTYTIEVQAYN 626
QY 420 VPSGQRFSLAIVH 433
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
627 VPSGQRFSLAIVN 640

Search completed: March 31, 2004, 16:04:32
Job time : 48.1304 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:59:39 ; Search time 13.9892 Seconds
(without alignments)
1597.947 Million cell updates/sec

Title: US-09-985-689A-5
Perfect score: 2261
Sequence: 1 NDVARGIVKADVAQNNGLY.....EVAQNVPSGQRSLAIVH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep:*
4: /cgn2_6/prodata/2/iaa/5B COMB.pep:*
5: /cgn2_6/prodata/2/iaa/6CTUS COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2245	99.3	433	4	US-09-104-623A-4
2	2245	99.3	433	4	US-09-019-532-4
3	2245	99.3	433	4	US-09-338-746-4
4	2245	99.3	635	2	US-08-873-479-43
5	2044	90.4	641	2	US-08-873-479-42
6	2002.5	88.6	639	4	US-09-509-814A-4
7	1999.5	88.4	640	4	US-09-509-814A-8
8	1998.5	88.4	640	4	US-09-509-814A-6
9	1952.5	86.4	639	4	US-09-509-814A-1
10	1952.5	86.4	640	4	US-09-509-814A-2
11	1514	67.0	345	4	US-08-512-251A-10
12	1514	67.0	345	4	US-09-515-150A-10
13	1514	67.0	345	4	US-09-196-281-13
14	443	19.6	659	3	US-08-894-818B-1
15	443	19.6	659	4	US-09-445-472-12
16	422.5	18.7	412	4	US-09-445-472-1
17	422.5	18.7	522	3	US-08-894-818B-3
18	422.5	18.7	522	4	US-09-445-472-4
19	422.5	18.7	654	3	US-08-894-818B-35
20	422.5	18.7	654	4	US-09-445-472-16
21	401.5	17.8	659	3	US-08-894-818B-5
22	343	15.2	520	4	US-09-000-016-7
23	343	15.2	520	4	US-09-514-340-7
24	343	15.2	734	3	US-09-000-016-4
25	343	15.2	734	4	US-09-514-340-4
26	343	15.2	823	3	US-09-000-016-2
27	343	15.2	823	4	US-09-514-340-2

28	307	13.6	903	1	US-08-750-532-1	Sequence 1, Appli
29	307	13.6	1398	1	US-08-750-532-9	Sequence 9, Appli
30	307	13.6	1398	3	US-08-894-818B-8	Sequence 8, Appli
31	307	13.6	1398	4	US-09-445-472-6	Sequence 6, Appli
32	280.5	12.4	237	1	US-08-750-532-18	Sequence 18, Appli
33	272	12.0	275	2	US-08-750-406A-1	Sequence 1, Appli
34	272	12.0	275	3	US-09-327-118-1	Sequence 1, Appli
35	269	11.9	269	1	US-08-431-387-5	Sequence 5, Appli
36	269	11.9	269	1	US-08-431-387-6	Sequence 6, Appli
37	269	11.9	269	1	US-08-322-677A-10	Sequence 10, Appli
38	269	11.9	269	1	US-08-322-676-10	Sequence 10, Appli
39	269	11.9	269	2	US-08-140-083A-10	Sequence 10, Appli
40	269	11.9	269	3	US-08-898-218-10	Sequence 10, Appli
41	269	11.9	269	3	US-08-848-793-10	Sequence 5, Appli
42	269	11.9	269	3	US-09-255-502-5	Sequence 3, Appli
43	269	11.9	269	3	US-09-024-532-3	Sequence 4, Appli
44	269	11.9	269	3	US-08-369-050-4	Sequence 1, Appli
45	269	11.9	269	3	US-08-090-207-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-104-623A-4
; Sequence 4, Application US/09104623A
; Patent No. 6303752
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Fatum, Tine Muxoll
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Roggen, Erwin Ludo
; TITLE OF INVENTION: A Modified Polypeptide
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 63037520 No. 6303752disk of No. 6303752th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/104,623A
; FILING DATE: 25-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol
; REGISTRATION NUMBER: 36,993
; REFERENCE/POCKET NUMBER: 5256.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-857-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; STRAIN: Bacillus sp. Y
; US-09-104-623A-4

Query Match 99.3%; Score 2245; DB 4; Length 433;
Best Local Similarity 99.1%; Pred. No. 1.6e-166;
Matches 429; Conservative 1; Indels 0; Gaps 0;
QY 1 NDVARGIVKADVAQNNGLYGGQVAVADTGLDTRNDSSHEAFRGKITALVALGRTN 60

Db 1 NDVARGIVKADVAQNNYGLYGGQQLVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Qy 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Db 61 NASDPNGHGTHTVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Qy 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Db 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Qy 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSLAPDSFW 240
Db 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSLAPDSFW 240
Qy 241 ANYSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
Db 241 ANYSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
Qy 301 SGDOGWRVTLDKSLNVAAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
Db 301 SGDOGWRVTLDKSLNVAAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
Qy 361 YTLVNDLIDLVTAPNGQKYVGNDFSYPDNNWDRNNVENVFINAPQSGTYTIEVOAYNV 420
Db 361 YTLVNDLIDLVTAPNGQKYVGNDFSYPDNNWDRNNVENVFINAPQSGTYTIEVOAYNV 420
Qy 421 PSGPQRFSLAIVH 433
Db 421 PSGPQRFSLAIVH 433

RESULT 2

US-09-019-532-4
; Sequence 4, Application US/09019532B
; Patent No. 6416756
; GENERAL INFORMATION:
; APPLICANT: Olsen, Arne Agerlin
; APPLICANT: Olsen, Annette
; TITLE OF INVENTION: A Modified Enzyme for Skin Care
; FILE REFERENCE: 4922.204-US
; CURRENT APPLICATION NUMBER: US/09/019.532B
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: 0038/97
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 0754/97
; EARLIER FILING DATE: 1997-06-25
; EARLIER APPLICATION NUMBER: 60/051,381
; EARLIER FILING DATE: 1997-07-07
; EARLIER APPLICATION NUMBER: PCT/DK98/00015
; EARLIER FILING DATE: 1998-01-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-019-532-4

Query Match 99.3%; Score 2245; DB 4; Length 433;
Best Local Similarity 99.1%; Pred. No. 1.6e-166;
Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 NDVARGIVKADVAQNNYGLYGGQQLVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQNNYGLYGGQQLVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Qy 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Db 61 NASDPNGHGTHTVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Qy 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180

Db 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Qy 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSLAPDSFW 240
Db 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSLAPDSFW 240
Qy 241 ANYSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
Db 241 ANYSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
Qy 301 SGDOGWRVTLDKSLNVAAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
Db 301 SGDOGWRVTLDKSLNVAAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360
Qy 361 YTLVNDLIDLVTAPNGQKYVGNDFSYPDNNWDRNNVENVFINAPQSGTYTIEVOAYNV 420
Db 361 YTLVNDLIDLVTAPNGQKYVGNDFSYPDNNWDRNNVENVFINAPQSGTYTIEVOAYNV 420
Qy 421 PSGPQRFSLAIVH 433
Db 421 PSGPQRFSLAIVH 433

RESULT 3

US-09-338-746-4
; Sequence 4, Application US/09338746
; Patent No. 6638526
; GENERAL INFORMATION:
; APPLICANT: Deussen, Heinz-Josef
; APPLICANT: Olsen, Arne A.
; APPLICANT: Fatum, Tine M.
; APPLICANT: Roggen, Erwin L.
; TITLE OF INVENTION: A Polypeptide-Polymer Conjugate
; FILE REFERENCE: 5619.200-US
; CURRENT APPLICATION NUMBER: US/09/338.746
; CURRENT FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: PA 1998 00809
; EARLIER FILING DATE: 1998-06-23
; EARLIER APPLICATION NUMBER: 60/091,461
; EARLIER FILING DATE: 1998-07-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-338-746-4

Query Match 99.3%; Score 2245; DB 4; Length 433;
Best Local Similarity 99.1%; Pred. No. 1.6e-166;
Matches 429; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 NDVARGIVKADVAQNNYGLYGGQQLVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db 1 NDVARGIVKADVAQNNYGLYGGQQLVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Qy 61 NANDPNHGHTHVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Db 61 NASDPNGHGTHTVAGSVLGNALNKGMAPOANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Qy 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Db 121 GARIHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAIT 180
Qy 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSLAPDSFW 240
Db 181 VGATENYRPSFGSLADPNHIAQFSSRGATRDGRIKPDVTAPGTFTLSARSSLAPDSFW 240
Qy 241 ANYSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
Db 241 ANYSKYAYMGGTSMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVGLGYP 300
Qy 301 SGDOGWRVTLDKSLNVAAYNEATATGOKATYSFOAQAGKPLKISLVWTDAPGSTTAS 360

Db 301 SDQGWGRVTLDKSLNVAIVNEATATATGQKATYSFQAQKPKLSLVWTDAPGSTAS 360
QY 361 YTLVNDLVLITAPNGQKYGNDFSYPYDNNWGDGRNNVNFINAPOSQGTITIEVQAYNV 420
Db 361 YTLVNDLVLITAPNGQKYGNDFSYPYDNNWGDGRNNVNFINAPOSQGTITIEVQAYNV 420
QY 421 PSGQRFSLAIVH 433
Db 421 PSGQRFSLAIVH 433

RESULT 4
US-08-873-479-43
; Sequence 43, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agtis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-873-479-43

Query Match 99.3%; Score 2245; DB 2; Length 635;
Best Local Similarity 99.1%; Pred. No. 2.7e-166;
Matches 423; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVQNNYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db 203 NDVARGIVKADVQNNYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 262

QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLFQAWNA 120
Db 263 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLFQAWNA 322

QY 121 GARHHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 180
Db 323 GARHHTNSWGPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAIT 382

QY 181 VGATENYRPSFGSLADPNPHIAQFSSRGATRGRIKPDVTAPGTFTLSARSSILAPSSFW 240
Db 383 VGATENYRPSFGSLADPNPHIAQFSSRGATRGRIKPDVTAPGTFTLSARSSILAPSSFW 442

QY 241 ANYNSKIYMGSTMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVLGLVP 300
Db 443 ANYNSKIYMGSTMATPIVAGNVAQLREHFINKRGITPKPSLIKAALIAGATDVLGLVP 502
QY 301 SGQGWGRVTLDKSLNVAIVNEATATATGQKATYSFQAQKPKLSLVWTDAPGSTAS 360
Db 503 SGQGWGRVTLDKSLNVAIVNEATATATGQKATYSFQAQKPKLSLVWTDAPGSTAS 562
QY 361 YTLVNDLVLITAPNGQKYGNDFSYPYDNNWGDGRNNVNFINAPOSQGTITIEVQAYNV 420
Db 563 YTLVNDLVLITAPNGQKYGNDFSYPYDNNWGDGRNNVNFINAPOSQGTITIEVQAYNV 622
QY 421 PSGQRFSLAIVH 433
Db 623 PSGQRFSLAIVH 635

RESULT 5
US-08-873-479-42
; Sequence 42, Application US/08873479
; Patent No. 5891701
; GENERAL INFORMATION:
; APPLICANT: Sloma, Alan
; APPLICANT: Lynne, Christianson
; TITLE OF INVENTION: Nucleic Acids Encoding A Polypeptide
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5891701o No. 5891701disk of No. 5891701th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 12-JUN-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Agtis, Cheryl H
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 5251.000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 641 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-873-479-42

Query Match 90.4%; Score 2044; DB 2; Length 641;
Best Local Similarity 89.1%; Pred. No. 1.1e-150;
Matches 386; Conservative 24; Mismatches 23; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVQNNYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db 209 NDVARGIVKADVQNNYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 268
QY 61 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLFQAWNA 120
Db 269 NANDPNHGHTHVAGSVLGNALNKGMAPQANLVFQSIIMDSGGGLGGLPSNLTLFQAWNA 328

QY 121 GARIHNSGAPVNGAYTANSQVDEYVENNDMTVLEAGNEGNSGTISAPGTAKNAIT 180
Db 329 GARIHNSGAPVNGAYTANSQVDEYVENNDMTVLEAGNEGNSGTISAPGTAKNAIT 388
QY 181 VGATENRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTPAGTIFLSARSLAPDSFW 240
Db 389 VGATENRPSFGSYADNINHVAQFSSRGATRGRIKPDVTPAGTIFLSARSLAPDSFW 448
QY 241 ANVNSKYAYMGTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGATDVGLGY 300
Db 449 ANHDSKYAYMGTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGATDVGLGY 508
QY 301 PSDQGWGRVTLDKSLNVAYNEATATGQKATYSFOAQKPLKISLVWTDAPGSTTA 360
Db 509 NGNQGWGRVTLDKSLNVAFVNETSPSTSQKATYSFTAQAGKPLKISLVWSDAPGSTTA 568
QY 361 YTLVNDLDLIVITAPNGQKYGNDPSYPYDNNNDGRNNVENFINAPOSGTITIEVQAYN 420
Db 569 LTLVNDLDLIVITAPNGTKYVNDFTAPYDNNNDGRNNVENFINAPOSGTITIEVQAYN 628
QY 421 PSQPQRFSLAIVH 433
Db 629 PVSQPTFSLAIVH 641

RESULT 6

US-09-509-814A-4
; Sequence 4, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:

; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JF98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 4
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.

US-09-509-814A-4

Query Match 88.6%; Score 2002.5; DB 4; Length 639;
Best Local Similarity 88.2%; Pred. No. 1.9e-147;
Matches 383; Conservative 26; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db 206 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 265
QY 61 NADPNHGHTHVAGSVLGN-ALNKGMAPQANLVFQSIMDSGGJGLGLPSNLTLFQAWN 119
Db 266 NADPNHGHTHVAGSVLGNATNKGMAPQANLVFQSIMDSGGJGLGLPSNLTLFQAFS 325
QY 120 AGARIHNSWGAPVNGAYTANSROVDYVRNNDMTVLEAGNEGNSGTISAPGTAKNAI 179
Db 326 AGARIHNSWGAAVNGAYTDSRVDDYVRKNDMTILFAAGNEPNEGTTISAPGTAKNAI 385
QY 180 TVGATENRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTPAGTIFLSARSLAPDSFW 239

Db 386 TVGATENRPSFGSYADNINHVAQFSSRGATKDGRIKPDVMAPGTIYLSARSLAPDSFW 445
QY 240 WANYNSKYAYMGTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGATDVGLGY 299
Db 446 WANHDSKYAYMGTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGADVGLGY 505
QY 300 PSDQGWGRVTLDKSLNVAYNEATATGQKATYSFOAQKPLKISLVWTDAPGSTTA 359
Db 506 PNGNQGWGRVTLDKSLNVAYNESSALSTSQKATYFTTAGKPLKISLVWSDAPASTTA 565
QY 360 SYTLVNDLDLIVITAPNGQKYGNDPSYPYDNNNDGRNNVENFINAPOSGTITIEVQAYN 419
Db 566 SVTLVNDLDLIVITAPNGTRYVNDPSAPPDNNNDGRNNVENFINPSQSGTITIEVQAYN 625
QY 420 VPSQPQRFSLAIVH 433
Db 626 VPVGPQNFSLAIVN 639

RESULT 7

US-09-509-814A-8
; Sequence 8, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:

; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUW
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-0PCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JF98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 8
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Bacillus sp.

US-09-509-814A-8

Query Match 88.4%; Score 1999.5; DB 4; Length 640;
Best Local Similarity 87.8%; Pred. No. 3.2e-147;
Matches 381; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNGYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
Db 207 NDVARGIVKADVAQSSYGLYGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 266
QY 61 NADPNHGHTHVAGSVLGN-ALNKGMAPQANLVFQSIMDSGGJGLGLPSNLTLFQAWN 119
Db 267 NADPNHGHTHVAGSVLGNSTNKGMAPQANLVFQSIMDSGGJGLGLPSNLTLFQAYS 326
QY 120 AGARIHNSWGAPVNGAYTANSROVDYVRNNDMTVLEAGNEGNSGTISAPGTAKNAI 179
Db 327 AGARIHNSWGAAVNGAYTDSRVDDYVRKNDMTILFAAGNEPNEGTTISAPGTAKNAI 386
QY 180 TVGATENRPSFGSLADNPNHIAQFSSRGATRGRIKPDVTPAGTIFLSARSLAPDSFW 239
Db 387 TVGATENRPSFGSYADNINHVAQFSSRGATKDGRIKPDVTPAGTIFLSARSLAPDSFW 446
QY 240 WANYNSKYAYMGTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGATDVGLGY 299
Db 447 WANHDSKYAYMGTSMATPIVAGNVAQLREHFVKNRGITPKPSLKAALIAGADVGLGY 506
QY 300 PSDQGWGRVTLDKSLNVAYNEATATGQKATYSFOAQKPLKISLVWTDAPGSTTA 359

```
Db 507 PGNQGWGRVLDLSDLVNAYVNESLSSTQKATVSFTATAGKPLKISLVWSDAPASTTA 566
QY 360 SYTLVNDLDLVITAFNGQKRYGNDSPYDNNWGRNNVNFVINAPOSQGYTIEVQAYN 419
Db 567 SVTLVNDLDLVITAFNGQYVYVNDFTSPYDNNWGRNNVNFVINAPOSQGYTIEVQAYN 626
QY 420 VPSGPQSFSLAIVH 433
Db 627 VPVGPQSFSLAIVN 640

RESULT 8
US-09-509-814A-6
; Sequence 6, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 6
; LENGTH: 840
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-509-814A-6

Query Match 88.4%; Score 1998.5; DB 4; Length 640;
Best Local Similarity 87.8%; Pred. No. 3 8e-147;
Matches 381; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNGYLGQSVVAVADTGLDTRNDSSMHEAPRGKITALYALGRTN 60
Db 207 NDVARGIVKADVAQSVGLYQSQIVAVADTGLDTRNDSSMHEAPRGKITALYALGRTN 266
QY 61 NANDPNHGHTHVAGSVLGN-ALNKGWAPQANLVFQSIMDSSGGLGSLPSNLATLFSQAWN 119
Db 267 NANDTNGHGHVAGSVLGNSTNGKWAQANLVFQSIMDSSGGLGSLPSNLATLFSQAYS 326
QY 120 AGARIHNSWGAPVNGAYTANSQVDEYVRNNDMTVLFAAGNEGPNSTISAPGTAKNAI 179
Db 327 AGARIHNSWGAAVNGAYTTDSRVDDYVRKNDWTILFAAGNEGPNSTISAPGTAKNAI 386
QY 180 TVGATENRPSFGSLADNPNHIAQFSRGATRCRKPDPVAPGTIFLSARSLAPDSF 239
Db 387 TVGATENRPSFGSLADNPNHIAQFSRGATRCRKPDPVAPGTIFLSARSLAPDSF 446
QY 240 WANVNSKYAVNGGTSMATPTVAGNVAQLRHHFKNRGITPKPSLIKAALIAGATDVGLGY 299
Db 447 WANHDSKYAVNGGTSMATPTVAGNVAQLRHHFKNRGITPKPSLIKAALIAGAADIGLY 506
QY 300 PSGDQGWGRVTLDKSLNVAVNBATALATQCKATYSFQAQAGKPLKISLVWTDAPGSTTA 359
Db 507 PGNQGWGRVTLDKSLNVAVNBATALATQCKATYSFQAQAGKPLKISLVWSDAPASTTA 566
QY 360 SYTLVNDLDLVITAPNGQKRYGNDSPYDNNWGRNNVNFVINAPOSQGYTIEVQAYN 419
Db 567 SVTLVNDLDLVITAPNGQYVYVNDFTSPYDNNWGRNNVNFVINAPOSQGYTIEVQAYN 626
```

```
QY 420 VPSGPQSFSLAIVH 433
Db 627 VPVGPQSFSLAIVN 640

RESULT 9
US-09-509-814A-1
; Sequence 1, Application US/09509814A
; Patent No. 6376227
; GENERAL INFORMATION:
; APPLICANT: TAKAIWA, MIKIO
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KUBOTA, HIROMI
; APPLICANT: HITOMI, JUN
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SHIKATA, SHITSUM
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: ALKALINE PROTEASE
; FILE REFERENCE: 0327-0832-OPCT
; CURRENT APPLICATION NUMBER: US/09/509,814A
; CURRENT FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: PCT/JP98/04528
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: JP 9-274570
; PRIOR FILING DATE: 1997-06-08
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Bacillus sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (23)..(23)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (29)..(29)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (32)..(32)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (46)..(46)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (53)..(53)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (70)..(70)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (89)..(89)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (102)..(102)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (105)..(105)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: (130)..(130)
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
```

LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (146)..(146)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (148)..(148)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (160)..(160)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (165)..(165)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (172)..(172)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (183)..(183)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (187)..(187)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (194)..(194)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (286)..(286)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (306)..(306)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (324)..(324)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (369)..(369)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (431)..(431)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (501)..(501)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (531)..(531)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (541)..(541)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (584)..(584)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (591)..(591)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (594)..(594)

OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (611)..(611)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc feature
LOCATION: (632)..(632)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-1
Query Match 86.4%; Score 1952.5; DB 4; Length 639;
Best Local Similarity 86.6%; Pred. No. 1.4e-143;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;
Qy 1 NDVARGIVKADVAQNNYGLYGQGVVAVADTGLDTGRNDSSMHEAFRGKITALLYALGRTN 60
Db 206 NDVARGIVKADVAQSSYGLYGQQIVAVADTGLDTGRNDSSMHEAFRGKITALLYALGRTN 265
Qy 61 NANDPNGHGHVAGSVLGN-ALNKGMAPQANLVFQSIMDSSGGLGGLPSNLTLPQAWN 119
Db 266 NANDTNGHGHVAGSVLGNKXNKGMAPQANLVFQSIMDSSGGLGGLPSNLTLPQAXS 325
Qy 120 AGARIHTNSGAPVNGAYTANSRQVDEYVRNDMTVLFAAGNEGPNSTGTSAPCTAKNAI 179
Db 326 AGARIHTNSGAAVNGAYTTDSRVDVYVRKNMTLFAAGNEXPNGGTISAGTAKNAI 385
Qy 180 TVGATENYRPSFGSLADNPNHIAQFSRGATROGRKIPDVAPGTFILSARSLAPDSSF 239
Db 386 TVGATENLRPSFGSYADNINHVAFSSRGPTKDGRIKPDVMAPTXILSARSLAPDSSF 445
Qy 240 WANYNSKYAYMGTSMATPIVAGNVAOLREHFKNRGITPKPSLIKAALITAGATDVLGY 299
Db 446 WANHDSKIAYMGTSMATPIVAGNVAOLREHFVKNRGITPKPSLKAALITAGADVLGY 505
Qy 300 PSGDQGWGRVTLDKSLNVAYVNEATATLQCKATYSFQAQAGKPLKISLVWTDAPGSTTA 359
Db 506 PNGNQGWGRVTLDKSLNVAYVNESSXLSTSQATYATATAGKPLKISLVWDAPASTTA 565
Qy 360 SYTLVNDLDELVITAPNGQKYGNDPFPYDNDNDGRNNVENFVNAPOSGCTYITIEVOAYN 419
Db 566 SVTLVNDLDELVITAPNGTXYVGNDFXXPXXXNWDGRNNVENFINPQSGTYYITIEVOAYN 625
Qy 420 VPSGQRFSLAIVH 433
Db 626 VPVGQXFLAIVN 639
RESULT 10
US-09-509-814A-2
Sequence 2, Application US/09509814A
Patent No. 6376227
GENERAL INFORMATION:
APPLICANT: TAKAIWA, MIKIO
APPLICANT: OKUDA, MITSUOSHI
APPLICANT: SAEKI, KATSUHIISA
APPLICANT: KUBOTA, HIROMI
APPLICANT: HITOMI, JUN
APPLICANT: KAGEYAMA, YASUSHI
APPLICANT: SHIKATA, SHITSUW
APPLICANT: NOMURA, MASAFUMI
TITLE OF INVENTION: ALKALINE PROTEASE
FILE REFERENCE: 0327-0832-0PCT
CURRENT APPLICATION NUMBER: US/09/509.814A
CURRENT FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: PCT/JP98/04528
PRIOR FILING DATE: 1998-10-07
PRIOR APPLICATION NUMBER: JP 9-274570
PRIOR FILING DATE: 1997-06-08

NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 640
TYPE: PRT
ORGANISM: Bacillus sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (24)..(24)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (30)..(30)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (33)..(33)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (47)..(47)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (48)..(48)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (54)..(54)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (71)..(71)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (75)..(75)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (90)..(90)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (103)..(103)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (106)..(106)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (129)..(129)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (131)..(131)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (132)..(132)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (133)..(133)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (134)..(134)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (147)..(147)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (149)..(149)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (161)..(161)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (166)..(166)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (173)..(173)
OTHER INFORMATION: Xaa is any amino acid

NAME/KEY: misc_feature
LOCATION: (184)..(184)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (188)..(188)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (189)..(189)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (190)..(190)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (195)..(195)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (287)..(287)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (307)..(307)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (325)..(325)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (370)..(370)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (432)..(432)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (502)..(502)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (532)..(532)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (542)..(542)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (585)..(585)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (592)..(592)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (593)..(593)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (595)..(595)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (596)..(596)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (597)..(597)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (612)..(612)
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: (633)..(633)
OTHER INFORMATION: Xaa is any amino acid
US-09-509-814A-2

Query Match 86.4%; Score 1952.5; DB 4; Length 640;
Best Local Similarity 86.6%; Pred. No. 1.4e-143;
Matches 376; Conservative 20; Mismatches 37; Indels 1; Gaps 1;

QY 1 NDVARGIVKADVAQNNGLYGOGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60
DB 207 NDVARGIVKADVAQSSYGLYGGQIVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 266

QY 61 NNDPNGHGHVAGSVLGN -ALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 119
Db 267 NNDPNGHGHVAGSVLGN -ALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 326
QY 120 AGARHTNSWGAPVNGAYTANSROVDEYVRNNDMTLFAAGNEGPNSTGTSAPGTAKNAI 179
Db 327 AGARHTNSWGAPVNGAYTANSROVDEYVRNNDMTLFAAGNEGPNSTGTSAPGTAKNAI 386
QY 180 TVGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 239
Db 387 TVGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 446
QY 240 WANYNSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLIKAALIAGATDVGLGYP 299
Db 447 WANDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLIKAALIAGATDVGLGYP 506
QY 300 PGGDQGWGRVTLDKSLNVAQNNFGLYQGGQIVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 359
Db 507 PGGDQGWGRVTLDKSLNVAQNNFGLYQGGQIVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 566
QY 360 SVTLVNDLVLITAPNGQKIVGNDPSYPYDNNWDGNNVNFVNAPOSGTIVIEVQAYN 419
Db 567 SVTLVNDLVLITAPNGQKIVGNDPSYPYDNNWDGNNVNFVNAPOSGTIVIEVQAYN 626
QY 420 VPSGPQRFSLAIVH 433
Db 627 VPSGPQRFSLAIVH 640

RESULT 11
US-09-512-251A-10
; Sequence 10, Application US/09512251A
; Patent No. 655335

GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5349.204-US
; CURRENT APPLICATION NUMBER: US/09/512,251A
; CURRENT FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Bacillus

Query Match 67.0%; Score 1514; DB 4; Length 345;

Best Local Similarity 90.2%; Pred. No. 7e-110;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNYGLYQGGQIVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
Db 29 NDVARGIVKADVAQNNYGLYQGGQIVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 88
QY 61 NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 120
Db 89 NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 148
QY 121 GARIHTNSWGAPVNGAYTANSROVDEYVRNNDMTLFAAGNEGPNSTGTSAPGTAKNAIT 180
Db 149 GARIHTNSWGAPVNGAYTANSROVDEYVRNNDMTLFAAGNEGPNSTGTSAPGTAKNAIT 208
QY 181 VGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 240
Db 209 VGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 268
QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLIKAALIAGATDVGLGYP 300
Db 269 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLIKAALIAGATDVGLGYP 328

QY 301 SGDQGWGRVTLDKSLNV 317
Db 329 NGNQGWRVTLDKSLNV 345

RESULT 12

US-09-515-150A-10
; Sequence 10, Application US/09515150A
; Patent No. 6558938
GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; APPLICANT: Andersen, Kim
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5348.204-US
; CURRENT APPLICATION NUMBER: US/09/515,150A
; CURRENT FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; TYPE: PRT
; ORGANISM: Bacillus

Query Match 67.0%; Score 1514; DB 4; Length 345;

Best Local Similarity 90.2%; Pred. No. 7e-110;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNYGLYQGGQIVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
Db 29 NDVARGIVKADVAQNNYGLYQGGQIVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 88
QY 61 NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 120
Db 89 NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWN 148
QY 121 GARIHTNSWGAPVNGAYTANSROVDEYVRNNDMTLFAAGNEGPNSTGTSAPGTAKNAIT 180
Db 149 GARIHTNSWGAPVNGAYTANSROVDEYVRNNDMTLFAAGNEGPNSTGTSAPGTAKNAIT 208
QY 181 VGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 240
Db 209 VGATENRPSFGSLADPNHIAQFSSRGATDRGRIKPDVTAPGTFTILSARSSSLAPDSSF 268
QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLIKAALIAGATDVGLGYP 300
Db 269 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFVKNRGVTPKPSLIKAALIAGATDVGLGYP 328

RESULT 13

US-09-196-281-13
; Sequence 13, Application US/09196281A
; Patent No. 6605458
GENERAL INFORMATION:
; APPLICANT: Hansen, Peter K.
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5435.200-US
; CURRENT APPLICATION NUMBER: US/09/196,281A
; CURRENT FILING DATE: 1998-11-19
; EARLIER APPLICATION NUMBER: 1332/97
; EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PastSeq for Windows Version 3.0
; SEQ ID NO 13

```
; LENGTH: 345
; TYPE: PR1
; ORGANISM: Bacillus
US-09-196-281-13

Query Match      67.0%; Score 1514; DB 4; Length 345;
Best Local Similarity 90.2%; Pred. No. 7e-110;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNGYLGQGVAVADTGLDTRNDSSMHEAFRGKITALLYALGRTN 60
Db 29 NDVARGIVKADVAQNNGYLGQGVAVADTGLDTRNDSSMHEAFRGKITALLYALGRTN 88
QY 61 NANDPNHGTHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPDLNLTLSQAWNA 120
Db 89 NANDPNHGTHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPDLNLTLSQAWNA 148
QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNDMTLFAAGNPGNSGTISAPGTAKNAIT 180
Db 149 GARIHTNSGAPVNGAYTANSRQVDEYVRNDMTLFAAGNPGNSGTISAPGTAKNAIT 208
QY 181 VGATENYRPSFGSLADPNHIAQFSRGATRDGRIPKDVTAFTGTFILSARSLAPDSFW 240
Db 209 VGATENYRPSFGSLADPNHIAQFSRGATRDGRIPKDVTAFTGTFILSARSLAPDSFW 268
QY 241 ANYNSKYAYMGTSMATPIVAGNVAQLREHFKNRGITPKSLIKAALTAGATDVGLGVP 300
Db 269 ANYNSKYAYMGTSMATPIVAGNVAQLREHFKNRGITPKSLIKAALTAGATDVGLGVP 328
QY 301 SGDSQGWGRVTLDKSLNV 317
Db 329 NGNQGWGRVTLDKSLNV 345

RESULT 14
US-08-894-818B-1
; Sequence 1, Application US/08894818B
; Patent No. 6261822
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: YAMAMOTO, Katsuhiko
; APPLICANT: MITTA, Masanori
; APPLICANT: ASADA, Kiyocho
; APPLICANT: TSUNASAWA, Susumu
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: United States of America
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,818B
; FILING DATE: 20-MAY-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP96/03253
; FILING DATE: 07-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 323285/1995
; FILING DATE: 12-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618

; REFERENCE/DOCKET NUMBER: TAKAKURA=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-894-818B-1

Query Match      19.6%; Score 443; DB 3; Length 659;
Best Local Similarity 31.2%; Pred. No. 2.8e-26;
Matches 137; Conservative 66; Mismatches 170; Indels 68; Gaps 16;

QY 8 VKADVAQNNGYLGQGVAVADTGLDTRNDSSMHEAFRGKITALLY-ALGRTNANDPN 66
Db 145 IGADTVNNSLGYDGSVWVAIVDTGIDAN-----HPDLKKGKVIWYDAVNGRSTPYDDQ 198
QY 67 GHGTHVAGSVLG-----NALNKGMAPQANLVFQSIM--DSSGGI GGLPDLNLTLSQAWNA 120
Db 199 GHGTHVAGSVLG-----NALNKGMAPQANLVFQSIM--DSSGGI GGLPDLNLTLSQAWNA 120
QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNDMTLFAAGNPGNSGTISAPGTAKNAIT 180
Db 259 GIRVINLSLSSQSSDGTDSLQAVNNADAGIWCVAAGNSGPNVTYVGSAAAASKVIT 318
QY 181 VGATENYRPSFGSLADPNHIAQFSRGATRDGRIPKDVTAFTGTFILSARSLAPDSFW 240
Db 319 VGA-----VDSNDNIASFSSRGPTADGRLEKPEVAVPGVDIAPRAS---GTSMG 364
QY 241 ANYNSKYAYMGTSMATPIVAGNVAQLREHFKNRGITP---KPSLIKAALTAGATDVG- 296
Db 365 TPINDYTKASGTSKATPHVSGVALILQ---AHPSTWPKVKTALITETADIVAPKEIAD 421
QY 297 LGYPSGQGWGRVTLDKSLNVAYVNEATATATGQ-----KATYSFQAQAGKPLKISLWVT 351
Db 422 IAY-----GAGRVNVYKA--IKYDDYAKLTFTGSVADKGSATHTFDVSGATFTVATLYWD 474
QY 352 DAGSTTASVTLVNDLDELITAPNGQKVGNDPFPYDNNWDRNNVNYFINAPQSGTY 411
Db 475 -----TGSSDIDILYLDPN-----GNEVDYSYATY---GFEKVGYNPTAGTW 515
QY 412 TIEVQAVNVPSGQPFSLAIV 432
Db 516 TVKVVSY---KGAANYQVDVV 533

RESULT 15
US-09-445-472-12
; Sequence 12, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMJO, Tomoko
; APPLICANT: ASADA, Kiyocho
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/09/445,472
; CURRENT FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PR1
; ORGANISM: Thermococcus celer
US-09-445-472-12
```

Query Match 19.6%; Score 443; DB 4; Length 659;
Best Local Similarity 31.1%; Pred.No. 2.8e-26;
Matches 137; Conservative 66; Mismatches 170; Indels 68; Gaps 16;

QY	8	VKADVAQNNYGLYGQGVAVADTGLDTRNDSSNHEAFRGKITALLY-ALGRTNNANDPN	66
Db	145	IGADTVMSNLGYDGSQGVVAIVDTGIDAN-----HPDLKGKVIQWYDAVNGRSTPYDDQ	198
QY	67	GHGTHVAGSVLG-----NALNKGMAQANLVFQSIM--DSSGGLGGLPSNLNTLFSQAWNA	120
Db	199	GHGTHVAGIVAGTGSVNSQYIGVAFGAKLVGVKLGADGSGSVSTIIAGVDVVQNKDKY	258
QY	121	GARIHTNSMGAPVNGAYTANSRQVDEYVVRNNDMTVLPAAAGNEGPNSTISAPGTAKNAIT	180
Db	259	GIRVINLSLGSQSSDGTDSLQAVNNAWDAGIVVCVAAGNSGPNNTYVGSPPAAASKVIT	318
QY	181	VGATENYRPSFGSLADNPNHIAQFSRGATRDGRKIPDVTAPGTFFLSARSSLAPDSFW	240
Db	319	VGA-----VDSNDNIASFSSRGPTADGRLKPEVAFPGVDIIAPRAS---GTSMG	364
QY	241	ANVNSKYAVMGTSMATPIVAGNVAQLREHFKNRGITP---KPSLIKAALIAGATDVG-	296
Db	365	TPINDYYTKASGTSVATPHVSGVGLILQ---AHPSTWTPDKVKYKTALITADIAPKEIAD	421
QY	297	LGYPESGDQGWGRVTLDKSLNVAYVNEATALATGQ-----KATYFQAQAGKPLKLSLWT	351
Db	422	IAY-----GAGRNVVYKA--IKYDDYAKLTFTGCVADKGSATHTFDVSGATFVTATLYWD	474
QY	352	DAPGSTTASVTLVNDLIDLITAPNGOKYVGNDFSPYPYDNNWDGRNNVENVFINAPQSGTY	411
Db	475	-----TGSSDIDLILYDPN-----GNEVDYSYAYY---GFEKVGYYNPTAGTW	515
QY	412	TIEVQAYNVPSGPFQFSLAIV	432
Db	516	TVKWSY---KGAANYQVDV	533

Search completed: March 31, 2004, 16:11:55
Job time : 14.9892 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 16:09:09 ; Search time 33.1412 Seconds
(without alignments)
3418.697 Million cell updates/sec

Title: US-09-985-689A-5
Perfect score: 2261
Sequence: 1 NDVARGIVKADVAQNNYGLY.....EVOQNVFSGQRFSLAIVH 433

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1065169 seqs, 261661801 residues

Total number of hits satisfying chosen parameters: 1065169

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubaa/US07_PUBCOMB.pep:
2: /cgn2_6/ptodata/1/pubaa/PTC_NEW_PUB.pep:
3: /cgn2_6/ptodata/1/pubaa/US05_NEW_PUB.pep:
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCOMB.pep:
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUB.pep:
6: /cgn2_6/ptodata/1/pubaa/PTUS_PUBCOMB.pep:
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUB.pep:
8: /cgn2_6/ptodata/1/pubaa/US09_PUBCOMB.pep:
9: /cgn2_6/ptodata/1/pubaa/US09A_PUBCOMB.pep:
10: /cgn2_6/ptodata/1/pubaa/US09B_PUBCOMB.pep:
11: /cgn2_6/ptodata/1/pubaa/US09C_PUBCOMB.pep:
12: /cgn2_6/ptodata/1/pubaa/US09A_NEW_PUB.pep:
13: /cgn2_6/ptodata/1/pubaa/US10A_PUBCOMB.pep:
14: /cgn2_6/ptodata/1/pubaa/US10B_PUBCOMB.pep:
15: /cgn2_6/ptodata/1/pubaa/US10C_PUBCOMB.pep:
16: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUB.pep:
17: /cgn2_6/ptodata/1/pubaa/US60_NEW_PUB.pep:
18: /cgn2_6/ptodata/1/pubaa/US60_PUBCOMB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2261	100.0	433	10 US-09-985-689A-5	Sequence 5, Appli
2	2251	99.6	433	10 US-09-985-689A-3	Sequence 3, Appli
3	2240	99.1	433	10 US-09-985-689A-4	Sequence 4, Appli
4	2044	90.4	433	10 US-09-985-689A-7	Sequence 7, Appli
5	2024	89.5	434	10 US-09-985-689A-6	Sequence 6, Appli
6	2010	88.9	434	10 US-09-985-689A-2	Sequence 2, Appli
7	1998.5	88.4	434	10 US-09-985-689A-1	Sequence 1, Appli
8	1998.5	88.4	434	15 US-10-385-662-2	Sequence 10, Appli
9	1514	67.0	345	14 US-10-336-324-10	Sequence 13, Appli
10	1514	67.0	345	14 US-10-403-105-13	Sequence 12, Appli
11	443	19.6	659	13 US-10-090-624-12	Sequence 1, Appli
12	422.5	18.7	412	13 US-10-090-624-1	Sequence 4, Appli
13	422.5	18.7	522	13 US-10-090-624-4	Sequence 16, Appli
14	422.5	18.7	654	13 US-10-090-624-16	Sequence 39, Appli
15	347.5	15.4	1079	14 US-10-112-488-39	

16	341.5	15.1	1237	14	US-10-314-657-4	Sequence 4, Appli
17	329	14.6	1139	14	US-10-156-761-10856	Sequence 10856, A
18	327	14.5	1208	14	US-10-156-761-13251	Sequence 13251, A
19	307	13.6	1398	13	US-10-090-624-6	Sequence 6, Appli
20	298.5	13.2	580	10	US-09-927-827-55	Sequence 55, Appli
21	295	13.0	539	15	US-10-084-848A-114	Sequence 114, Appli
22	295	13.0	19725	15	US-10-084-848A-4	Sequence 4, Appli
23	284.5	12.6	1101	14	US-10-156-761-12934	Sequence 12934, A
24	276.5	12.2	368	12	US-10-344-231-3	Sequence 3, Appli
25	276	12.2	271	10	US-09-813-408-2	Sequence 2, Appli
26	271	12.0	271	14	US-10-242-549-56	Sequence 56, Appli
27	270	11.9	271	14	US-10-242-549-54	Sequence 54, Appli
28	270	11.9	271	14	US-10-242-549-60	Sequence 60, Appli
29	269	11.9	269	8	US-08-322-678-10	Sequence 10, Appli
30	269	11.9	269	9	US-09-837-235-16	Sequence 16, Appli
31	269	11.9	269	9	US-09-060-854B-6	Sequence 6, Appli
32	269	11.9	269	9	US-09-975-139-1	Sequence 1, Appli
33	269	11.9	269	9	US-09-976-414-8	Sequence 8, Appli
34	269	11.9	269	10	US-09-736-116-49	Sequence 49, Appli
35	269	11.9	269	13	US-10-075-907-1	Sequence 1, Appli
36	269	11.9	269	13	US-10-075-895-1	Sequence 1, Appli
37	269	11.9	269	14	US-10-033-325-6	Sequence 6, Appli
38	269	11.9	269	14	US-10-209-812-3	Sequence 3, Appli
39	269	11.9	269	14	US-10-104-693-4	Sequence 4, Appli
40	269	11.9	269	14	US-10-228-572-6	Sequence 6, Appli
41	269	11.9	269	14	US-10-324-152-5	Sequence 5, Appli
42	269	11.9	269	14	US-10-242-549-5	Sequence 5, Appli
43	269	11.9	269	14	US-10-313-853-1	Sequence 1, Appli
44	269	11.9	269	14	US-10-336-324-4	Sequence 4, Appli
45	269	11.9	269	14	US-10-336-324-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1

US-09-985-689A-5
; Sequence 5, Application US/09985689A
; Publication No. US2003022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-5

Query Match 100.0%; Score 2261; DB 10; Length 433;
Best Local Similarity 100.0%; Pred. No. 1.1e-191;
Matches 433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NDVARGIVKADVAQNNYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60

Db 1 NDVARGIVKADVAQNNYGLYGGQGVAVADTGLDTGRNDSMHEAFRGKITALYALGRTN 60

QY 61 NANDPNHGTHVAGSVLGNALNKGMAPQANLVFQSIQNDSSGGIGGLPNTLTFQAWNA 120


```

Db      61  NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY      121  GARIHTNSGAPVNGAYTANSROVDEYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
Db      121  GARIHTNSGAPVNGAYTANSROVDEYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
QY      181  VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
Db      181  VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
QY      241  ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
Db      241  ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
QY      301  SGDOGWRVTLDKSLNVAYVNEATATGOKATYSFOAQAGKPLKISLWMTDAPGSTTAS 360
Db      301  SGDOGWRVTLDKSLNVAYVNEATATGOKATYSFOAQAGKPLKISLWMTDAPGSTTAS 360
QY      361  YTLVNDLVLVITAPNGQKYVGNDFSYPDNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
Db      361  YTLVNDLVLVITAPNGQKYVGNDFSYPDNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
QY      421  PSGPQRFSLAIVH 433
Db      421  PSGPQRFSLAIVH 433

```

RESULT 2

```

US-09-985-689A-3
; Sequence 3, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-3

```

```

Query Match      99.6%; Score 2251; DB 10; Length 433;
Best Local Similarity 99.3%; Pred. No. 8.7e-191;
Matches 430; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1  NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMSHEAPRGKITALYALGRTN 60
Db      1  NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMSHEAPRGKITALYALGRTN 60
QY      61  NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Db      61  NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY      121  GARIHTNSGAPVNGAYTANSROVDEYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
Db      121  GARIHTNSGAPVNGAYTANSROVDEYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180

```

```

QY      181  VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
Db      181  VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
QY      241  ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
Db      241  ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
QY      301  SGDOGWRVTLDKSLNVAYVNEATATGOKATYSFOAQAGKPLKISLWMTDAPGSTTAS 360
Db      301  SGDOGWRVTLDKSLNVAYVNEATATGOKATYSFOAQAGKPLKISLWMTDAPGSTTAS 360
QY      361  YTLVNDLVLVITAPNGQKYVGNDFSYPDNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
Db      361  YTLVNDLVLVITAPNGQKYVGNDFSYPDNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
QY      421  PSGPQRFSLAIVH 433
Db      421  PSGPQRFSLAIVH 433

RESULT 3
US-09-985-689A-4
; Sequence 4, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-4

```

```

Query Match      99.1%; Score 2240; DB 10; Length 433;
Best Local Similarity 98.8%; Pred. No. 8.2e-190;
Matches 428; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      1  NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMSHEAPRGKITALYALGRTN 60
Db      1  NDVARGIVKADVAQNNYGLYGQGVAVADTGLDTGRNDSMSHEAPRGKITALYALGRTN 60
QY      61  NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
Db      61  NNDPNGHGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGGLPSNLNTLFSQAWNA 120
QY      121  GARIHTNSGAPVNGAYTANSROVDEYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
Db      121  GARIHTNSGAPVNGAYTANSROVDEYVRNNDMTVLFAAGNEGPNSTGTTISAPGTAKNAIT 180
QY      181  VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
Db      181  VGATENYRPSFGSLADPNPHIAQFSSRGATRDGRKIPDVTAPGTFILSARSSSLAPDSSFW 240
QY      241  ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300
Db      241  ANYNSKYAYMGTSMTPIVAGNVAQLREHFINKRGITPKPSLIKAALITAGATDVGLGYP 300

```

```
QY 301 SDGQGRVTLDKSLNVAVNEATATLCTCKATYSFQAQAGKPLKISLWTDAPGSTTAS 360
Db 301 NGDQGRVTLNKSUNVAVNEATATLCTCKATYSFQAQAGKPLKISLWTDAPGSTTAS 360
QY 361 YTLVNDLDLIVITAPNGKTVGNDFSYYPYDNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
Db 361 YTLVNDLDLIVITAPNGKTVGNDFSYYPYDNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
QY 421 PSQPQFSLAIVH 433
Db 421 PSQPQFSLAIVH 433

RESULT 4
US-09-985-689A-7
; Sequence 7, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-7

Query Match 90.4%; Score 2044; DB 10; Length 433;
Best Local Similarity 89.1%; Pred. No. 1.9e-172; Indels 0; Gaps 0;
Matches 386; Conservative 24; Mismatches 23;

QY 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
Db 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNALNKGWAPQANLVFQSTMDSSGGGLGPSNLNTLFSQAWN 120
Db 61 NANDPNHGHTHVAGSVLGNATNKGWAPQANLVFQSTMDSSGGGLGPSNLNTLFSQAYSA 120
QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVNNDMTILFAAGNEGPNSGTISAPGTAKNAI 180
Db 121 GARIHTNSGAPVNGAYTTDSRNVDDYVKNDMTILFAAGNEGPGSGTISAPGTAKNAI 180
QY 181 VGATENYRPSGLADNPHIAQFSSRGATRGRIKPDVTAQGTILSARSLAPDSSF 240
Db 181 VGATENLRFPSGSYADNINHVAQFSSRGPTRGRIKPDVMAFGTILSARSLAPDSSF 240
QY 241 ANYNSKYAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALITAGATDVGLGY 300
Db 241 ANHDSKYAYMGGTSMATPIVAGNVAQLREHFKVNGRVTGPKPSLLKXALJAGAADVGLGFP 300
QY 301 SDGQGRVTLDKSLNVAVNEATATLCTCKATYSFQAQAGKPLKISLWTDAPGSTTAS 360
Db 301 NGDQGRVTLDKSLNVAFVNETSPLSTQKATYSFTAQAGKPLKISLWSDAPGSTTAS 360
QY 361 YTLVNDLDLIVITAPNGKTVGNDFSYYPYDNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
Db 361 YTLVNDLDLIVITAPNGKTVGNDFSYYPYDNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
```

```
Db 361 LTLVNDLDLIVITAPNGTKYVGNDFTAIPYDNNWDGRNNVENVFINAPQSGTYTIEVQAYNV 420
QY 421 PSQPQFSLAIVH 433
Db 421 PVSFQTFSLAIVH 433

RESULT 5
US-09-985-689A-6
; Sequence 6, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483US0
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-6

Query Match 89.5%; Score 2024.5; DB 10; Length 434;
Best Local Similarity 88.5%; Pred. No. 1e-170; Indels 1; Gaps 1;
Matches 384; Conservative 28; Mismatches 21;

QY 1 NDVARGIVKADVAQNNGYLYGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
Db 1 NDVARGIVKADVAQNSYLYGQGVVAVADTGLDTCGRNDSMHEAFRGKITAYALGRTN 60
QY 61 NANDPNHGHTHVAGSVLGNALNKGWAPQANLVFQSTMDSSGGGLGPSNLNTLFSQAWN 119
Db 61 NANDPNHGHTHVAGSVLGNATNKGWAPQANLVFQSTMDSSGGGLGPSNVSTLFSQAYS 120
QY 120 AGARIHTNSGAPVNGAYTANSRQVDEYVNNDMTILFAAGNEGPNSGTISAPGTAKNAI 179
Db 121 AGARIHTNSGAPVNGAYTTDSRNVDDYVKNDMAVLFAAGNEGPNGGTISAPGTAKNAI 180
QY 180 TVGATENYRPSGLADNPHIAQFSSRGATRGRIKPDVTAQGTILSARSLAPDSSF 239
Db 181 TVGATENLRFPSGSYADNINHVAQFSSRGPTRGRIKPDVMAFGTILSARSLAPDSSF 240
QY 240 WANYNSKYAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLIKAALITAGATDVGLGY 299
Db 241 WANHDSKYAYMGGTSMATPIVAGNVAQLREHFIKNGRITPKPSLLKXALJAGATDGLGY 300
QY 300 PSQDQGRVTLDKSLNVAVNEATATLCTCKATYSFQAQAGKPLKISLWTDAPGSTTAS 359
Db 301 PSQDQGRVTLDKSLNVAFVNETSPLSTQKATYSFTAQSGKPLKISLWSDAPASTSA 360
QY 360 SYTLVNDLDLIVITAPNGKTVGNDFSYYPYDNNWDGRNNVENVFINAPQSGTYTIEVQAYN 419
Db 361 SVTLVNDLDLIVITAPNGTKYVGNDFTAIPYDNNWDGRNNVENVFINAPQSGTYTIEVQAYN 420
QY 420 VPSQPQFSLAIVH 433
Db 421 VPQGPQAFSLAIVN 434
```

RESULT 6

US-09-985-689A-2
; Sequence 2, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI
; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-2

Query Match 88.9%; Score 2010.5; DB 10; Length 434;

Best Local Similarity 88.5%; Pred. No. 1.8e-169;

Matches 384; Conservative 26; Mismatches 23; Indels 1; Gaps 1;

QY	1	NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGRNDSSMHEAPRGKITALYALGRN	60
DB	1	NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGRNDSSMHEAPRGKITALYALGRN	60
QY	61	NANDPNGHGHVAGSVLGN-ALNKGMAPQANLVFQSIMDSSGGGLGSLNLTLSQAWN	119
DB	61	NANDTNGHGHVAGSVLGNATNKGMAPQANLVFQSIMDSSGGGLGSLNLTLSQAYS	120
QY	120	AGARIHNSGAPNGAYTANSROVDYVNRNDMTVLFAAGNEGPNSTISAPGTAKNAI	179
DB	121	AGARIHNSGAAVNGAYTTDSRVDDYVRKNDMTILFAAGNEGPNSTISAPGTAKNAI	180
QY	180	TVGATENRPSFGSLADPNPHIAQFSSRGATRDGRIPKDPVTAPGTFILSARSLAPDSSF	239
DB	181	TVGATENLRPSFGSADYNHVAQFSSRGPTKGRIPKDPVMAFGTVILSARSLAPDSSF	240
QY	240	WANYNSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGY	299
DB	241	WANHDSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLLKAALIAGAADVGLGY	300
QY	300	PSGQGWGRVTLDKSLNVAVNEATATGQKATYSFQAQAGKPLKISLVWTDAPGSTTA	359
DB	301	PNGNQGWGRVTLDKSLNVAVYNESSLSSTQKATYFTATAGKPLKISLVWSDAPASTTA	360
QY	360	SYTLVNDLDELVTAPNGQKTVGNDFSYPYDNNWGRNNVNFINAPQSGTYTIEVOAYN	419
DB	361	SVTLVNDLDELVTAPNGTRYVGNDFSPYDNNWGRNNVNFINAPQSGTYTIEVOAYN	420
QY	420	VPSGPQRFSLAIVH 433	
DB	421	VPVGPQNFSLAIVN 434	

RESULT 7

US-09-985-689A-1
; Sequence 1, Application US/09985689A
; Publication No. US20030022351A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, YUJI
; APPLICANT: OGAWA, AKINORI

; APPLICANT: KAGEYAMA, YASUSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: ARAKI, HIROYUKI
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SAEKI, KATSUHIISA
; TITLE OF INVENTION: Alkaline proteases
; FILE REFERENCE: 215483USO
; CURRENT APPLICATION NUMBER: US/09/985,689A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: JP P2000-355166
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: JP P2001-114048
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-09-985-689A-1

Query Match 88.4%; Score 1998.5; DB 10; Length 434;

Best Local Similarity 87.8%; Pred. No. 2.1e-168;

Matches 381; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

QY	1	NDVARGIVKADVAQNNYGLYGQGVAVADTGLDGRNDSSMHEAPRGKITALYALGRN	60
DB	1	NDVARGIVKADVAQSSYGLYGQGVAVADTGLDGRNDSSMHEAPRGKITALYALGRN	60
QY	61	NANDPNGHGHVAGSVLGN-ALNKGMAPQANLVFQSIMDSSGGGLGSLNLTLSQAWN	119
DB	61	NANDTNGHGHVAGSVLGNATNKGMAPQANLVFQSIMDSSGGGLGSLNLTLSQAYS	120
QY	120	AGARIHNSGAPNGAYTANSROVDYVNRNDMTVLFAAGNEGPNSTISAPGTAKNAI	179
DB	121	AGARIHNSGAAVNGAYTTDSRVDDYVRKNDMTILFAAGNEGPNSTISAPGTAKNAI	180
QY	180	TVGATENRPSFGSLADPNPHIAQFSSRGATRDGRIPKDPVTAPGTFILSARSLAPDSSF	239
DB	181	TVGATENLRPSFGSADYNHVAQFSSRGPTKGRIPKDPVMAFGTVILSARSLAPDSSF	240
QY	240	WANYNSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAGATDVGLGY	299
DB	241	WANHDSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLLKAALIAGAADVGLGY	300
QY	300	PSGQGWGRVTLDKSLNVAVNEATATGQKATYSFQAQAGKPLKISLVWTDAPGSTTA	359
DB	301	PNGNQGWGRVTLDKSLNVAVYNESSLSSTQKATYFTATAGKPLKISLVWSDAPASTTA	360
QY	360	SYTLVNDLDELVTAPNGQKTVGNDFSYPYDNNWGRNNVNFINAPQSGTYTIEVOAYN	419
DB	361	SVTLVNDLDELVTAPNGTRYVGNDFSPYDNNWGRNNVNFINAPQSGTYTIEVOAYN	420
QY	420	VPSGPQRFSLAIVH 433	
DB	421	VPVGPQNFSLAIVN 434	

RESULT 8

US-10-385-662-2
; Sequence 2, Application US/10385662
; Publication No. US20040002432A1
; GENERAL INFORMATION:
; APPLICANT: OKUDA, MITSUYOSHI
; APPLICANT: SATO, TSUYOSHI
; APPLICANT: SAITO, KAZUHIRO
; APPLICANT: SUMITOMO, NOBUYUKI
; APPLICANT: IZAWA, YOSHIKUMI
; APPLICANT: SAEKI, KATSUHIISA
; APPLICANT: KOBAYASHI, TOHRU
; APPLICANT: NOMURA, MASAFUMI
; TITLE OF INVENTION: Alkaline protease

```
; FILE REFERENCE: 234938US0
; CURRENT APPLICATION NUMBER: US/10/385,662
; PRIOR FILING DATE: 2003-03-12
; PRIOR APPLICATION NUMBER: JP 2002-081428
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: JP 2002-165987
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: JP 2002-304230
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: JP 2002-304231
; PRIOR FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Bacillus sp. KSM-KP43
US-10-385-662-2

Query Match      88.4%; Score 1998.5; DB 15; Length 434;
Best Local Similarity 87.8%; Pred. No. 2.1e-168;
Matches 381; Conservative 28; Mismatches 24; Indels 1; Gaps 1;

QY      1 NDVARGIVKADVAQNNGYLYGGQVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Db      1 NDVARGIVKADVAQSSYGLYGGQIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
QY      61 NANDPNHGHTHVAGSVLGN-ALNKGWAPQANLVFQSIMDSSGGLGLPSNLTLFSQAWN 119
Db      61 NANDPNHGHTHVAGSVLGNSTNGKMAPQANLVFQSIMDSSGGLGLPSNLTLFSQAYS 120
QY      120 AGARIHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTGISAPGTAKNAI 179
Db      121 AGARIHTNSWGAAGVNGAYTTDSRNVDDYVRKNDMTILFAAGNEGPNSTGISAPGTAKNAI 180
QY      180 TVGATENYRPSFGSLADNPENHIAQFSSRGATDRGRIKPDVTAPGTFILSARSSSLAPDSF 239
Db      181 TVGATENLRPSFGSYADNHNHVAQFSSRGPTDGRGRIKPDVWAPGTFILSARSSSLAPDSF 240
QY      240 WANYNSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAAGATDVGLGY 299
Db      241 WANHDSKYAYMGTSMTATPIVAGNVAQLREHFIKNGRITPKPSLIKAALIAAGADIGLY 300
QY      300 PSGDQGWGRVTLDKSLUNVAYNEATALATQKATYSFOQAQKPLKISLVWTDAPGSTTA 359
Db      301 PNGNQGWGRVTLDKSLUNVAYNESSLSQKATYSFATAGKPLKISLVMSDAPASTTA 360
QY      360 SVTLVNDLVLITAPNGQKVGNDPSYVDNKNWGRNNVNFVINAPOSGTITIEVQAYN 419
Db      361 SVTLVNDLVLITAPNGTQVGNDFSTPNDNWDGRNNVNFVINAPOSGTITIEVQAYN 420
QY      420 VPSGPQRFSLAIYH 433
Db      421 VPVGPQTFSIAIYN 434

RESULT 9
US-10-336-324-10
; Sequence 10, Application US/10336324
; Publication No. US20030176304A1
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants and Compositions
; FILE REFERENCE: 5349.204-US
; CURRENT APPLICATION NUMBER: US/10/336,324
; PRIOR FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: US/09/512,251A
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 10
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-10-336-324-10

Query Match      67.0%; Score 1514; DB 14; Length 345;
Best Local Similarity 90.2%; Pred. No. 1.3e-125;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY      1 NDVARGIVKADVAQNNGYLYGGQVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Db      29 NDVARGIVKADVAQNNGYLYGGQIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 88
QY      61 NANDPNHGHTHVAGSVLGNALNKGWAPQANLVFQSIMDSSGGLGLPSNLTLFSQAWN 120
Db      89 NANDPNHGHTHVAGSVLGNATNKGWAPQANLVFQSIMDSSGGLGLPANLQTLFSQAYS 148
QY      121 GARLHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTGISAPGTAKNAI 180

; SEQ ID NO 13
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-10-403-105-13
; Sequence 13, Application US/10403105
; Publication No. US20030180933A1
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter K.
; APPLICANT: Bauditz, Peter
; APPLICANT: Mikkelsen, Frank
; TITLE OF INVENTION: Protease Variants And Compositions
; FILE REFERENCE: 5435.200-US
; CURRENT APPLICATION NUMBER: US/10/403,105
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US/09/196,281A
; PRIOR FILING DATE: 1998-11-19
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 1332/97
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-11-21
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Bacillus
US-10-403-105-13

Query Match      67.0%; Score 1514; DB 14; Length 345;
Best Local Similarity 90.2%; Pred. No. 1.3e-125;
Matches 286; Conservative 18; Mismatches 13; Indels 0; Gaps 0;

QY      1 NDVARGIVKADVAQNNGYLYGGQVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 60
Db      29 NDVARGIVKADVAQNNGYLYGGQIVAVADTGLDTRNDSSMHEAFRGKITALYALGRTN 88
QY      61 NANDPNHGHTHVAGSVLGNALNKGWAPQANLVFQSIMDSSGGLGLPSNLTLFSQAWN 120
Db      89 NANDPNHGHTHVAGSVLGNATNKGWAPQANLVFQSIMDSSGGLGLPANLQTLFSQAYS 148
QY      121 GARLHTNSWGAPVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSTGISAPGTAKNAI 180
```

Db 149 GARIHTNSGAPVNGAYTTDSRNVDDYVKNNDMTILFAAGNEGPGSGTISAPGTAKNAIT 208
QY 181 VGATENYRSPFSGSLADPNPHIAQFSSRGATRDGRIPKDPVTAGTFTILGARSSLAPDSSFW 240
Db 209 VGATENLRFSGSYADNINHVAQFSSRGTRDGRIPKDPVMAFGTYILSARSSLAPDSSFW 268
QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFHFKNRIKGTIPKPSLIKAALIAGATDVLGYP 300
Db 269 ANHDSKYAYMGTSMTATPIVAGNVAQLREHFHFKNRIKGTIPKPSLIKAALIAGAADVLGFP 328
QY 301 SGDQGWGRVTLDKSLNV 317
Db 329 NGNQGWGRVTLDKSLNV 345

RESULT 11
US-10-090-624-12
; Sequence 12, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090,624
; CURRENT FILING DATE: 2002-03-06
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Thermococcus celer
US-10-090-624-12

Query Match 19.6%; Score 443; DB 13; Length 659;
Best Local Similarity 31.1%; Pred. No. 2,8e-30;
Matches 137; Conservative 66; Mismatches 170; Indels 69; Gaps 16;

QY 8 VKADVAQNNGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALLYALGRTNNAN-----DNGHG 69
Db 145 IGADTVNLSLGYDGSQVVAIVDTGIDAN-----HPDLKGVIGWYDAVNGRSTPYDDQ 198
QY 67 GHGTHVAGSVLG-----NALNKGMAPOANLVFOSIM--DSSGGLGGLPSNLTLFQAANA 120
Db 199 GHGTHVAGSVGNSQVIGVAPGAKLVGVKVLGADGSGSVSTIIAGVDVWVQNKOKY 258
QY 121 GARIHTNSGAPVNGAYTANSRQVDEYVRNNDMTILFAAGNEGPGSGTISAPGTAKNAIT 180
Db 259 GIRVINLSLSSQSSDGTSLSQAVNANDAGIVVCVAAAGNSGPNITYTVGSPAASKVIT 318
QY 181 VGATENYRSPFSGSLADPNPHIAQFSSRGATRDGRIPKDPVTAGTFTILSARSSLAPDSSFW 240
Db 319 VGA-----VDSNDNIASFSSRGPTADGRLKPEVAPGVPIIAPRAS---GTSMG 364
QY 241 ANYNSKYAYMGTSMTATPIVAGNVAQLREHFHFKNRIKGTIP---KPSLIKAALIAGATDVG- 296
Db 365 TPINDYTKRASGTSMTATPHVSGVALILQ---AHPSWTPDKVKLTALLETADIVAPKEIAD 421
QY 297 LGVPSDQGWGRVTLDKSLNVAVNEATLATGQ-----KATYSFOAQAGKPLKISLVMT 351
Db 422 IAY-----GAGRVNVYKA--IKYDDYAKLTFTGSAVDKGSATHTFDVSGATEVATLYWD 474
QY 352 DAPGSTTASVTLVNDLVLITAFNGQKYVGNDFSYPYDNNWDRNNVNFVINAPOSQTY 411
Db 475 -----TGSSDIDLVLIDFN-----GNEVDISYTAAY---GFEKVGYYNPTAGTW 515

QY 412 TIEVOANYVPSPQRFSLAIV 432
Db 516 TVKVVSY---KGAANYQVDVW 533

RESULT 12
US-10-090-624-1
; Sequence 1, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
; APPLICANT: TAKAKURA, Hikaru
; APPLICANT: MORISHITA, Mio
; APPLICANT: SHIMOJO, Tomoko
; APPLICANT: ASADA, Kiyozo
; APPLICANT: KATO, Ikunoshin
; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
; FILE REFERENCE: TAKAKURA=6
; CURRENT APPLICATION NUMBER: US/10/090,624
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: 09/445,472
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 151969/1997
; PRIOR FILING DATE: 1997-06-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
US-10-090-624-1

Query Match 18.7%; Score 422.5; DB 13; Length 412;
Best Local Similarity 31.4%; Pred. No. 9.1e-29;
Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;

QY 16 NYGLYGQGVVAVADTGLDTRNDSSMHEAFRGKITALLYALGRTNNAN-----DNGHG 69
Db 20 NLGYDGSGITIGIDTGID-----ASHPDLOGKV-----IGWDFVNGRSYPYDDHGHG 68
QY 70 THVAGSVLG-----NALNKGMAPOANLVFOSIM--DSSGGLGGLPSNLTLFQAANA 122
Db 69 THVASIAAGTGAASNGKYKMAFGAKLAGIKVLGADGSGSISTIIKGVWAVDNKDKYGI 128
QY 123 RIHTNSGAPVNGAYTANSRQVDEYVRNNDMTILFAAGNEGPGSGTISAPGTAKNAITVG 182
Db 129 KVINLSLSSQSSDGTSLSQAVNAAWDAGLVVVAAGNSGPNKYTIGSPAASKVITVG 188
QY 183 ATENYRSPFSGSLADPNPHIAQFSSRGATRDGRIPKDPVTAGTFTILSARSSLAPDSSFWAN 242
Db 189 AVDKY-----DVITSFSSRGPTADGRLKPEVAPGVPIIAPRAS---GTSMGQP 234
QY 243 YNSKYAYMGTSMTATPIVAGNVAQLREHFHFKNRIKGTIP---KPSLIKAALIAGATDVG-LG 298
Db 235 INDYTTAAPTGSMTATPHVAGIAALLQ---AHPSWTPDKVKLTALLETADIVKPEIADIA 291
QY 299 YPSDQGWGRVTLDKSLNVAVNEATLATGKA-----TYSFOAQAGKPLKISLVMTDA 353
Db 292 Y-----GAGRVNVYKAIN--YDNYAKLVFTGYVANGSQTHQFVIGSAFVATILWDNA 344
QY 354 PGSTTASVTLVNDLVLITAFNGQKYVGNDFSYPYDNNWDRNNVNFVINAPOSQTYTI 413
Db 345 N-----SDDLVLVYDPNGQV---DYSY-----TAYYGFEEKVGYYNPTDGTWTI 385
QY 414 EVQANYVPSPQRFSLAIV 432
Db 386 KWSY---SGSANYQVDVW 401

RESULT 13
US-10-090-624-4
; Sequence 4, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:

; APPLICANT: TAKAKURA, Hikaru
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: SHIMOJO, Tomoko
 ; APPLICANT: ASADA, Kiyozo
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
 ; FILE REFERENCE: TAKAKURA=6
 ; CURRENT APPLICATION NUMBER: US/10/090,624
 ; CURRENT FILING DATE: 2002-03-06
 ; PRIOR APPLICATION NUMBER: 09/445,472
 ; PRIOR FILING DATE: 1999-12-06
 ; PRIOR APPLICATION NUMBER: 151969/1997
 ; PRIOR FILING DATE: 1997-06-10
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 522
 ; TYPE: PRT
 ; ORGANISM: Pyrococcus furiosus
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (428)..(428)
 ; OTHER INFORMATION: Xaa at position 428 is Gly or Val.
 US-10-090-624-4

Query Match 18.7%; Score 422.5; DB 13; Length 522;
 Best Local Similarity 31.4%; Pred. No. 1.3e-28;
 Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;
 QY 16 NYGLYGQGVAVADTGLDTGRNDSSMHEAFRGKITALYALGRTNAN-----DPNGHG 69
 Db 20 NLGYDGSGITIGIDTID-----ASHPLQGV-----IGWDFVNGRSYPYDDHGHG 68
 QY 70 THVAGSVLG-----NALNKGMAQANLVFQSIM--DSSGGLGGLPSNLTLFSQAWNAGA 122
 Db 69 THVASIAAGTGAASNGKYKGMAPCAKLAGIKVLGADGSGSISTIIKGVEAVDNKDKYGI 128
 QY 123 RIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAITVG 182
 Db 129 KVINLSLGSSQSDGTDALSQAVNAWDAGLVVVAAGSGPNKYIIGSPAAASKVITVG 188
 QY 183 ATENYRPSFGLADNPNHIAQFSRGATRDGRIPKDVPTAPGTFILSARSLAPDSSFWAN 242
 Db 189 AVDKY-----DVIITFSRSGPTADGRLEKPEVAVPAGNWIIAARAS---GTSMGOP 234
 QY 243 YNSKYAVMGTSMTPTVAGNVLAQLREHFIKNGRITP---KPSLIKAALIAAGATDVG-LG 298
 Db 235 INDYITAAPGTSMATPHVAGIAALLQ---AHPSTWPKVKKTALIEADIVKDEIADIA 291
 QY 299 YPSGDQGWRTLDKSLNVAVNEATALTGOKA-----TYSFQAQAGKPLKISLVWTD 353
 Db 292 Y-----GAGRVNAYKAIN--YDNYAKLVFTGYVANKGSQTHQFVISCASFVTATLYWDNA 344
 QY 354 PGSTTASYTLVNDLVLITAPNGQKYVGNDFSPYDNNWGRNVENVFINAPQSGTYTI 413
 Db 345 N-----SDLDLYLDPNGNQV---DYSY-----TAYYGFKEKGVYNTDGTWTI 385
 QY 414 EQVAYNVPSGPFPSLAIV 432
 Db 386 KWSY---SGSANYQVDVW 401

RESULT 14
 US-10-090-624-16
 ; Sequence 16, Application US/10090624
 ; Publication No. US20020132335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TAKAKURA, Hikaru
 ; APPLICANT: MORISHITA, Mio
 ; APPLICANT: SHIMOJO, Tomoko
 ; APPLICANT: ASADA, Kiyozo
 ; APPLICANT: KATO, Ikunoshin
 ; TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE

; FILE REFERENCE: TAKAKURA=6
 ; CURRENT APPLICATION NUMBER: US/10/090,624
 ; CURRENT FILING DATE: 2002-03-06
 ; PRIOR APPLICATION NUMBER: 09/445,472
 ; PRIOR FILING DATE: 1999-12-06
 ; PRIOR APPLICATION NUMBER: 151969/1997
 ; PRIOR FILING DATE: 1997-06-10
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 16
 ; LENGTH: 654
 ; TYPE: PRT
 ; ORGANISM: Pyrococcus furiosus
 US-10-090-624-16
 Query Match 18.7%; Score 422.5; DB 13; Length 654;
 Best Local Similarity 31.4%; Pred. No. 1.8e-28;
 Matches 138; Conservative 55; Mismatches 167; Indels 79; Gaps 17;
 QY 16 NYGLYGQGVAVADTGLDTGRNDSSMHEAFRGKITALYALGRTNAN-----DPNGHG 69
 Db 152 NLGYDGSGITIGIDTID-----ASHPLQGV-----IGWDFVNGRSYPYDDHGHG 200
 QY 70 THVAGSVLG-----NALNKGMAQANLVFQSIM--DSSGGLGGLPSNLTLFSQAWNAGA 122
 Db 201 THVASIAAGTGAASNGKYKGMAPCAKLAGIKVLGADGSGSISTIIKGVEAVDNKDKYGI 260
 QY 123 RIHTNSGAPVNGAYTANSRQVDEYVRNNDMTVLFAGNEGPNSTISAPGTAKNAITVG 182
 Db 261 KVINLSLGSSQSDGTDALSQAVNAWDAGLVVVAAGSGPNKYIIGSPAAASKVITVG 320
 QY 183 ATENYRPSFGLADNPNHIAQFSRGATRDGRIPKDVPTAPGTFILSARSLAPDSSFWAN 242
 Db 321 AVDKY-----DVIITFSRSGPTADGRLEKPEVAVPAGNWIIAARAS---GTSMGOP 365
 QY 243 YNSKYAVMGTSMTPTVAGNVLAQLREHFIKNGRITP---KPSLIKAALIAAGATDVG-LG 298
 Db 367 INDYITAAPGTSMATPHVAGIAALLQ---AHPSTWPKVKKTALIEADIVKDEIADIA 423
 QY 299 YPSGDQGWRTLDKSLNVAVNEATALTGOKA-----TYSFQAQAGKPLKISLVWTD 353
 Db 424 Y-----GAGRVNAYKAIN--YDNYAKLVFTGYVANKGSQTHQFVISCASFVTATLYWDNA 476
 QY 354 PGSTTASYTLVNDLVLITAPNGQKYVGNDFSPYDNNWGRNVENVFINAPQSGTYTI 413
 Db 477 N-----SDLDLYLDPNGNQV---DYSY-----TAYYGFKEKGVYNTDGTWTI 517
 QY 414 EQVAYNVPSGPFPSLAIV 432
 Db 518 KWSY---SGSANYQVDVW 533
 RESULT 15
 US-10-112-488-39
 ; Sequence 39, Application US/10112488
 ; Publication No. US20030082746A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIKUCHI, Yoshiaki
 ; APPLICANT: DATE, Masayo
 ; APPLICANT: YOKOYAMA, Keiichi
 ; APPLICANT: YOKOYAMA, Keiichi
 ; APPLICANT: MATSUI, Hiroshi
 ; TITLE OF INVENTION: PROCESS FOR PRODUCING TRANSLUTAMINASE
 ; FILE REFERENCE: 219286USOCONT
 ; CURRENT APPLICATION NUMBER: US/10/112,488
 ; CURRENT FILING DATE: 2002-04-01
 ; PRIOR APPLICATION NUMBER: PCT/JF00/06780
 ; PRIOR FILING DATE: 2000-09-29
 ; PRIOR APPLICATION NUMBER: JP2000-280098
 ; PRIOR FILING DATE: 2000-06-28
 ; PRIOR APPLICATION NUMBER: JP11-280098
 ; PRIOR FILING DATE: 1999-09-30
 ; NUMBER OF SEQ ID NOS: 70

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 39
; LENGTH: 1079
; TYPE: PRT
; ORGANISM: Streptomyces albogriseolus
US-10-112-488-39

Query Match          15.4%; Score 347.5; DB 14; Length 1079;
Best Local Similarity 29.7%; Pred. No. 1.7e-21;
Matches 130; Conservative 52; Mismatches 181; Indels 75; Gaps 16;

QY 18 GLYGOQVAVADTGLDGTGRNDSMEAFRGKITALYALGRNTNANDPNHGHTHVAGSVL 77
Db 189 GYDGKGVKIAVLDTGVD-----ATHPDLKGQVTSKNTSAPTGDVVGHGTHVASIAA 242
QY 78 GNALN-----KGNAPQANLVFQSIMDSSGGLGSLPSNLNTLFSQAWNA--GARIHTNSWG 130
Db 243 GTCAQSKGYKGVAPGAKILNGKVLDDA---GFGDDSGILAGMEWAAQAQADI VNMSLG 298
QY 131 A-----PVNGAYTANSRQVDEYVRNNDMTVLFAAGNEGPNSGTISAPGTAKNAITVGA 183
Db 299 GMDTPETDPLEAA-----VDKLSAEKGILFAIAAGNEGPOS--IGSPGSADSALTVGA 349
QY 184 TENYRFSFGLADNPNIHQFSSRG--ATRDGRIKPDVTAPGTFFILSARSSLAPDSSFWAN 242
Db 350 -----VDDKDLADFSSTGPRLGDAVKPDLTAPGVDITAAASAKGNDIAKEVGE 398
QY 243 YNSKYAYMGTSMATPIVAGNVAQLREHFINKRGITP--KPSLIKAALIAGATDVGLGYP 300
Db 399 KPAGYMTISGTSNATPHVAGAAALLKQOH-----PENKYAELKGALTASTKDG--XYT 449
QY 301 SGDOGWGRVTLDKSLNVAVYNEATALATG-----QKATYSTQAQAGKPLKIS 347
Db 450 PFEQGSGRVQVQDKAITQTIVIAEFVSLSGVQWPHADDXEVTKKLTYNLGTEDVTCLKT 509
QY 348 LVMTDAPG-STTASYTLVNDLIVITAPNGQKYVGNDFSPYDNNMWDGNNVNFINAP 406
Db 510 STATGPKGKAAPAGFTLGLASTLTVPA-NGTASVDVTADTRLGGAVDGTYSAYVWATGAG 568
QY 407 QS-----GTYTIEVQAYNV 420
Db 569 QSVRTAAAVEREVESYNV 586
```

Search completed: March 31, 2004, 16:34:01
Job time : 34.1412 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2004, 15:58:38 ; Search time 11.3246 Seconds
(without alignments)
3677.911 Million cell updates/sec

Title: US-09-985-689A-5
Perfect score: 2261
Sequence: 1 NDVARGIVKADVAQNNGLY.....EVQAYNVPSGQRFSLAIVH 433

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	501	22.2	1743	2	multidrug resistan
2	463.5	20.5	1905	2	multidrug resistan
3	344	15.2	444	2	intracellular alka
4	315.5	14.0	442	2	intracellular alka
5	312	13.8	806	2	microbial serine p
6	305	13.5	1398	2	pycolysin (EC 3.4
7	302.5	13.4	799	2	subtilisin-type pr
8	298.5	13.2	580	2	serine proteinase
9	285	12.6	1345	2	surface layer-asso
10	272.5	12.1	715	2	alkaline serine pr
11	269	11.9	380	2	high-alkaline seri
12	265	11.7	524	1	alkaline proteinas
13	263.5	11.7	420	1	subtilisin (EC 3.4
14	256.5	11.3	419	1	subtilisin (EC 3.4
15	256.5	11.3	1331	2	probable surface 1
16	255	11.3	401	2	serine proteinase
17	255	11.3	757	2	subtilisin-type pr
18	251.5	11.1	627	2	serine proteinase,
19	250	11.1	488	2	cell wall-associat
20	250	11.1	894	2	streptococcal Csa
21	244.5	10.8	1167	1	high-alkaline seri
22	243.5	10.8	378	2	subtilisin (EC 3.4
23	243	10.7	382	2	subtilisin (EC 3.4
24	242	10.7	379	1	SUBSCL
25	241.5	10.7	321	1	S27501
26	241	10.7	513	1	aqualysin (EC 3.4
27	238	10.5	384	2	alkaline proteinas
28	234.5	10.4	402	1	alkaline proteinas
29	232.5	10.3	319	2	microbial serine p

30	229.5	10.2	519	2	S71451	halolysin R4 (EC 3
31	228	10.1	1052	2	T17093	intraluminal subti
32	227.5	10.1	381	1	SUBSS	subtilisin (EC 3.4
33	227.5	10.1	381	2	JQ1487	subtilisin (EC 3.4
34	227.5	10.1	381	2	JH0778	subtilisin (EC 3.4
35	227.5	10.1	1374	2	D72593	hypothetical prote
36	226.5	10.0	381	1	SUBSI	subtilisin (EC 3.4
37	226.5	10.0	905	2	F82734	serine proteinase
38	226	10.0	613	2	S75976	hypothetical prote
39	225.5	10.0	1433	1	A36734	bacillopeptidase F
40	225	10.0	382	1	SUBSN	subtilisin (EC 3.4
41	224.5	9.9	272	2	A23624	subtilisin (EC 3.4
42	224	9.9	274	1	SUBSD	subtilisin (EC 3.4
43	224	9.9	601	2	JC4576	serine proteinase
44	223.5	9.9	279	1	SUMVTV	thermitase (EC 3.4
45	223	9.9	275	2	JC1085	subtilisin (EC 3.4

ALIGNMENTS

RESULT 1

T18279

multidrug resistance transport protein - slime mold (Dictyostelium discoideum)
C/Species: Dictyostelium discoideum

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T18279

R/Shaulsky, G.; Loomis, W.F.

submitted to the EMBL Data Library, June 1996

A/Reference number: Z18855

A/Accession: T18279

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1743 <SHA>

A/Cross-references: EMBL:U60086; NID:G1399914; PID:G1399915; PIDN:AAB03331.1

C/Genetics:

A/Gene: tagC

Query Match 22.2%; Score 501; DB 2; Length 1743;

Best local similarity 27.7%; Pred. No. 2.5e-24;

Matches 163; Conservative 77; Mismatches 158; Indels 190; Gaps 24;

Qy	19	LYGGOVAVADTGLDGR---	NDS-----SMHEAFRGKITALYALGRNTNANDPNGH	68
Db	314	LRKGQLSLADTGLGSHCFSDSKYPIPLSNVNLNR-KVVTYITTTSDSDSKVDGH	372	
Qy	69	GTHVAGSVLG-----NALN----	KGMAPQANLVFQSIMDSSGGGLGSLPSNLTILFSQAWNA	120
Db	373	GTHICGSAAGTPESSVNISSPSGLATDAKIAFFDLASGSSSLTP-PSDLKQLYQPLYDA	431	
Qy	121	GARIHNSWGA-----PVNGAYTANSQVDYV-RNNDMTVLPAAGNEGPNSTIS--APG	173	
Db	432	GARVHCDSWGSVSEGTGYSSTDSITASIDFLFTHPDFILRAAGN---NEQYLSLLTQS	498	
Qy	174	TAKNAITVGGATE-----NYRPS-----FQS-----	193	
Db	489	TAKNVIIVGAHQITHENYLTDGPNYNYQSVVDINQELICDFDSRYCNVTTAQCLESNA	548	
Qy	194	-----LADNPHIAQFSSRGATRGRIKPDVTPAGTITLSA	229	
Db	549	TTGLASCCPTLLRKSRVIDAANTOPLYNENNICSFSGKGTGDRMKPALVAPEVITSA	608	
Qy	230	RSSILA-----PSSSEWANYNSKYAYMGTSMTPTVAGNVAQLREH-----FIK	273	
Db	609	RSNGANTTDCGGDSL-PNTNALLA-ISGTSMTATSPAAAATTLRQYLVVDGYVPTGSIVE	666	
Qy	274	NRGTTPPSIUKAALIAGA-----TDVGLGVPSGD-----QMGWVTLTD	312	
Db	667	SNKLOPTGSLIKALMINNAQLNGTFLQITSSSITYPNSQNVFENFAGASLVQGWGAIMRS	726	
Qy	313	KSLNVAVNEAT-----ALATGQKATYSF-----	336	
Db	727	NLHVANNNNNNKNTSDGITKFDGIGGLDLRLVKPNQWKESLSLQNTSYCTYKPS	786	

A;Residues: 1-442 <KUN>
 A;Cross-references: GB:Z99113; GB:AL009126; NID:g2634090; PIDN:CAB13610.1; PID:el183385;
 A;Experimental source: strain 168
 C;Genetics:
 A;Gene: aprX
 C;Superfamily: subtilisin homology
 F:146-398/Domain: subtilisin homology <SBT>

Query Match 14.0%; Score 315.5; DB 2; Length 442;
 Best Local Similarity 28.6%; Pred. No. 4.2e-13;
 Matches 103; Conservative 44; Mismatches 102; Indels 111; Gaps 13;
 QY 11 DVAQNNYGLYGQGVAVADTGLDGRNDSSMHEAFRGKITALVALGRTNAN-----D 64
 DB 136 EYVRNGQTLTGKGVAVVDTGI-----YHPDLEGR-----IGFADVMNQKTEPYD 183
 QY 65 PNGHGHVAGSVLGNALN-----KGMAPQANLVFQSIMDSSG----- 101
 DB 184 DNGHGHCHAGDVASSGSSGQYRGAPEANLIGVKVLKQSGTLADIISGVWCIOYN 243
 QY 102 -----GLGG-----LPSNLTFLSQAWNAGARIHTNSWGAPVNGAYTAN 140
 DB 244 EDNPEPIDIMSGSLGDDALRDYDHEQEDPLVRAVEEWSAG----- 284
 QY 141 SRQVDYVRNNDMTVLFAAGNPGNSGTISAPGTAKNALITVGATENYRPSFGSLADNPNH 200
 DB 285 -----IVVCVAAGNSGPDQITIASPEVSEKVIITVGALDNN-----NTASSDDT 328
 QY 201 IAOFSRGATRGRIKPDYAPGTFTLSARSLAPDSSF-----WANYNSKYAVNGGTSM 255
 DB 329 VASFSRGTPTVYKKEPDIAPGVNIIISLRS---PNSYIDKLQKSRVGSQYFTMSGTSM 385
 QY 256 APTIVAGVNAQUREHIKNGRITPRLIKAAIAGATDVGILYPSGQCGWGRVTLDKSL 315
 DB 386 APTICAGIAALILQ---QNPDLTPDE--VK-ELLKNGTKDKWDEPNFYAGAVNAENSV 439

RESULT 5

A41341
 Microbial serine proteinase (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis
 C;Species: Bacillus subtilis
 C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 20-Jun-2000
 C;Accession: A41341; B41341; S39700; D69730
 R;Sloma, A.; Ruffo Jr., G.A.; Thériault, K.A.; Dwyer, M.; Wilson, S.W.; Pero, J.
 J. Bacteriol. 173, 6889-6895, 1991
 A;Title: Cloning and characterization of the gene for an additional extracellular serine
 A;Reference number: A41341; MUID:92041574; PMID:1938892
 A;Accession: A41341
 A;Molecule type: DNA
 A;Residues: 1-806 <SLO>
 A;Cross-references: GB:M76590; NID:G143819; PIDN:AAA22881.1; PID:G143820
 A;Accession: B41341
 A;Molecule type: protein
 A;Residues: 161-195 <SL2>
 R;Glaser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu,
 A.; Rapoport, G.; Danchin, A.
 Mol. Microbiol. 10, 371-384, 1993
 A;Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region fr
 A;Reference number: S39655; MUID:95020537; PMID:7934828
 A;Accession: S39700
 A;Molecule type: DNA
 A;Residues: 1-806 <GLA>
 A;Cross-references: EMBL:X73124; NID:G413923; PIDN:CAA51601.1; PID:G580871
 R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, C.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Kosono, S.; Hullo, M.F.
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 J.

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, R.; Sekiguchi, J.; Sekowska, A.; Sero
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
 A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A;Reference number: A69580; MUID:98044033; PMID:9884377
 A;Accession: D69730
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-806 <KUN>
 A;Cross-references: GB:Z99123; GB:AL009126; NID:g2636240; PIDN:CAB15835.1; PID:g2636344
 A;Experimental source: strain 168
 C;Comment: The amino terminal sequence of the mature protein and a molecular weight of
 C;Genetics:
 A;Gene: vpr
 A;Start codon: TTG
 C;Superfamily: microbial serine proteinase vpr; subtilisin homology
 C;Keywords: hydrolase; serine proteinase
 F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-160/Domain: propeptide #status predicted <PRO>
 F:180-548/Domain: subtilisin homology #status atypical <SBT>

Query Match 13.8%; Score 312; DB 2; Length 806;
 Best Local Similarity 23.4%; Pred. No. 1.6e-12;
 Matches 129; Conservative 56; Mismatches 156; Indels 210; Gaps 17;
 QY 18 GLYGQGVAVADTGLDTCR-----NDSSMHEAFRGKITALVALGRTN 60
 DB 177 GYTKGKIKVAIIDTGVVEYHNPDLKKNFGQYGVDFVNDYDKETFG----- 224
 QY 61 NAMDPN-----HGHVAGSVLGNALNKGMAPQANLVFQSIMDSSGGLGLPSNLTLFSQ 116
 DB 225 ---DPRCEATDHGTHVAGTVAANGTIKGVAPDATTLLAYRVLPGG--SGTTENVIAGVER 279
 QY 117 ANAGARIHTNSWGAPVNGAYTANSQVDYVRNNDMTVLFAAGNPGNSGTISAPGTAK 176
 DB 280 AVQGDADWNLSLGNLNPDPWATSTALD--WAMSEGVAVTNSGNSGPNQWTVSGPSTSR 338
 QY 177 NAITVGATE---NYRPSFGSL----- 194
 DB 339 EATSVGATQLPLNEYAVTFGSSYSAKVMGYNKEDDVKALNNKEVELVEAGIGEAQDFEGK 398
 QY 195 -----ADNP----- 198
 DB 399 DLTKGVAWVRKGSIAFYDKADNKAQGAIGAVVYNNLSGSEIANVPQMSVPTTKLSLEDG 458
 QY 199 -----NHIAQFSRGATRD--GRIKPDVTAFTPTLSARSS 232
 DB 459 EKLVSALKAGETKTTFKLTYSKALGEQVADFSSRGPVMDTWMIKPDISAFGVNIVSIPT 518
 QY 233 LAPDSSFWANYNSKYAVNGGTSMATPIVAGNVAQLREHFTKNRGITPKPSL--IKAAALIA 290
 DB 519 HDPDHPY--GYGSKQ---GTSMAHPHAGAVAVIKQ-----AKPKWSEVQIKAAIMN 565
 QY 291 GATDV---GLGYPSGQCGWGRVTLDKSLNAVYVNEATATATGQKATY--SFQAGAKPLK 345
 DB 566 TAVTLKDSGDGVYVPHNAQGAQ-----SARIMNAIKADSLVSPGSYSGYTFLEKNGNETK 619
 QY 346 ISLVMTDAPGSTTASYTLVNDLIDLIVTAPNQKYVGNDFSYPYDNNWGDGNNVENVFINA 405
 DB 620 NETFTIENQSSIRKSYTL-----EYSPNGSGISTSGTSRVRVIPA 658
 QY 406 PQSGTYTIEVQ 416
 DB 659 HQTGKATAKVK 669

RESULT 6

T28159
 Pyrolysins (EC 3.4.-.-) - Pyrococcus furiosus
 C;Species: Pyrococcus furiosus
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 24-Oct-2000
 C;Accession: T28159